

# **COTRANSCRIPTIONAL PROCESSING OF HISTONE PRE-mRNA**

Kelly Daniel Sullivan

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Biochemistry and Biophysics

Chapel Hill  
2008

Approved by:

Advisor: Dr William F. Marzluff

Reader: Dr Howard M. Fried

Reader: Dr Jean G. Cook

Reader: Dr. Steve Rogers

Reader: Brian D. Strahl

© 2009

Kelly Daniel Sullivan

ALL RIGHTS RESERVED

## Abstract

KELLY DANIEL SULLIVAN: Cotranscriptional Processing of Histone pre-mRNA

(Under the direction of Dr. William F. Marzluff)

Metazoan histone mRNAs are unique in that they lack introns and poly(A) tails; thus these mRNAs require a single endonucleolytic cleavage for maturation. Cleavage of histone mRNA is guided by two cis-elements, a stemloop (SL) 5' of the cleavage site and a 3' purine-rich histone downstream element (HDE). Binding of the stemloop binding protein to the SL and basepairing of the U7 snRNA component of the U7 snRNP recruit a cleavage factor that contains the endonuclease CPSF73 which cleaves the pre-mRNA. I show here that a subset of poly(A) factors are required for histone pre-mRNA processing in *Drosophila melanogaster*. I present evidence that CPSF73, CPSF100, and Symplekin are required for histone pre-mRNA processing and that they form a stable complex. RNAi depletion of these factors results in misprocessing and polyadenylation of endogenous histone H2A mRNA. Knockdown of CPSF73, CPSF100 or Symplekin results in codepletion of the other factors. Coimmunoprecipitation experiments show that these proteins interact with one another and with the histone specific factors SLBP and Lsm11. I present chromatin immunoprecipitation data demonstrating that CPSF73 and Symplekin associate with both histone and poly(A) genes at similar levels in vivo, but that CstF50, a poly(A) specific factor associates with poly(A) genes at a much higher

level. Experiments combining RNAi and ChIP reveal that knockdown of SLBP results in recruitment of additional components of the poly(A) apparatus (CstF50) to histone genes. I also present ChIP data which show that knockdown of factors (CPSF160, CstF64) that codeplete proteins required for both poly(A) and histone pre-mRNA processing (CPSF73, Symplekin) specifically deplete the pool of these factors associated with poly(A) genes, consistent with two distinct processing complexes. I also present evidence in mammalian cells that CPSF73 is the exonuclease which degrades the 5' downstream cleavage product potentially resulting in transcription termination. Finally, I describe a role for the histone pre-mRNA processing factor SLBP in nuclear export of the mature mRNA. A detailed molecular analysis of the histone mRNA in cells with extremely low levels of SLBP reveals that the major lesion in histone mRNA metabolism is retention of mature, properly processed mRNA in the nucleus.

## **Acknowledgements**

First and foremost I would like to thank my parents, Dan and Kathy, who have supported me in whatever endeavors I have chosen to undertake throughout my life. I would also like to thank my brother Shane, sister-in-law Charise, and sister Megan for supporting me.

The next person I would like to thank is my mentor, Bill Marzluff. Bill is a one of a kind mentor whose enthusiasm for science knows no bounds. When things got difficult for me, I could always count on an encouraging word from Bill to put me back on track. Bill taught me to think like a scientist, but perhaps more importantly he taught me to enjoy science.

I would like to thank the members of the Marzluff lab, past and present, for countless hours of scientific discussion as well as encouragement. In particular, I would like to thank Tom Mullen and Eric Wagner. These two members of the Marzluff lab were not only wonderful mentors to me, they were also wonderful friends.

I would like to acknowledge my committee: Jean Cook, Howard Fried, Steve Rogers and Brian Strahl, whose patience and guidance allowed me to complete this dissertation.

Finally, I would like to thank the many friends I was blessed enough to know during my time in Chapel Hill; I am eternally grateful to you all.

## Table of Contents

List of Figures.....	viii
----------------------	------

List of Abbreviations.....	x
----------------------------	---

### Chapter

I.	Introduction.....	1
	Histones and Chromatin Structure.....	1
	Histone Gene Organization and Histone Locus Bodies.....	2
	Histone Gene Expression.....	3
	Replication-dependent Histone mRNAs.....	4
	Mammalian Histone pre-mRNA 3' End Formation.....	5
	Histone pre-mRNA Processing in <i>Drosophila</i> .....	6
	Translation and Degradation of Histone mRNAs.....	7
	Cleavage and Polyadenylation of pre-mRNAs.....	8
	Cotranscriptional Processing of pre-mRNAs.....	11
	Transcription Termination.....	14
	Summary.....	15
II.	A Subset of Polyadenylation Factors is Required for Cotranscriptional Histone pre-mRNA Processing in <i>Drosophila</i> .....	31
	Introduction.....	31
	Materials and Methods.....	35

	Results.....	38
	Discussion.....	54
III.	CPSF73 is both the endonuclease and a processive 5' to 3' exonuclease in histone pre-mRNA processing.....	83
	Introduction.....	83
	Materials and Methods.....	85
	Results.....	89
	Discussion.....	98
IV.	SLBP Knockdown Results Nuclear Retention of Histone mRNA.....	115
	Introduction.....	115
	Materials and Methods.....	118
	Results.....	123
	Discussion.....	136
V.	Summary and Conclusions.....	159
	Introduction.....	159
	A subset of cleavage and polyadenylation factors forms a “holo-cleavage factor” active in 3' end formation of all mRNAs.....	160
	Stable complexes containing polyadenylation factors in <i>Drosophila</i> .....	161
	Histone 3' end formation occurs cotranscriptionally.....	164
	Transcription termination of histone genes does not require Xrn2.....	167
	SLBP functions in export of histone mRNA.....	169
	Final Conclusions.....	171
VI.	References.....	173

## List of Figures

Figure 1. Histone mRNA and the cell cycle.....	18
Figure 2. Schematic of histone pre-mRNA processing.....	20
Figure 3. Conservation of histone pre-mRNA processing machinery between flies and humans.....	22
Figure 4. Global view of histone mRNA metabolism in mammalian cells.....	24
Figure 5. Schematic of mammalian cleavage and polyadenylation machinery.....	26
Figure 6. Cotranscriptional pre-mRNA processing.....	28
Figure 7. Transcriptional termination models.....	30
Figure 8. Knockdown of pre-mRNA processing factors results in misprocessed histone mRNA.....	64
Figure 9. The core cleavage factor interacts with histone specific processing factors.....	66
Figure 10. Codepletion is not due to decreased mRNA levels.....	68
Figure 11. Interactions between the cleavage factor and histone specific RNA- processing factors are not RNA-dependent.....	70
Figure 12. RNA processing factors are recruited to histone genes in vivo.....	72
Figure 13. SLBP associates with histone genes in mammalian cells.....	74



Figure 14. Loss of a histone specific processing factor results in recruitment of additional poly(A) factors to histone genes.....	76
Figure 15. Knockdown of poly(A) specific factors does not affect recruitment of core cleavage factor to histone genes.....	78
Figure 16. 3' end processing factors associate with RNA Pol II.....	80
Figure 17. Model of histone pre-mRNA processing.....	82
Figure 18. The U7-dependent degradation of the downstream cleavage product generated during 3' end processing of histone pre-mRNAs.....	107
Figure 19. Degradation of the 3' extDCP RNA containing a single radioactive phosphate downstream of the HDE.....	109
Figure 20. Switching between the 5' exonuclease and endonuclease modes.....	111
Figure 21. Xrn2 is not required for degradation of HIS2H2AA downstream cleavage product in vivo.....	114
Figure 22. SLBP knockdown cells are viable but grow at a slower rate as a result of various S-phase defects.....	146
Figure 23. A minimal SLBP construct rescues the cell-cycle phenotype associated with SLBP knockdown.....	148
Figure 24. Histone mRNA levels are reduced in SLBP knockdown cells and a portion of the mRNA is misprocessed and polyadenylated.....	150
Figure 25. Histone protein levels are decreased in SLBP knockdown cells and normal histone mRNA degradation in response to HU is impaired.....	152
Figure 26. Histone mRNA in SLBP knockdown cells is retained in the nucleus.....	154
Figure 27. SLBP promotes mRNA export.....	156
Figure 28. Model of SLBP-dependent histone mRNA export.....	158

### **List of Abbreviations**

CB	cajal body
CPSF	cleavage and polyadenylation specificity factor
CstF	cleavage stimulatory factor
CTD	C-terminal domain of RNA polymerase II
HDE	histone downstream element
HLB	histone locus body
SL	stemloop
SLBP	stemloop binding protein
snRNA	small nuclear RNA
snRNP	small nuclear ribonucleoprotein

## **Chapter 1**

### **Introduction**

#### **Histones and Chromatin Structure**

The eukaryotic genome is organized into nucleoprotein complexes termed nucleosomes. These units are composed of an octamer of positively charged histone proteins and approximately 147 bp of DNA (Luger et al., 1997). Each nucleosomal octamer consists of a core of two H3-H4 dimers, flanked on each side by an H2A-H2B dimer. The core histone proteins have structured C-termini which interact with DNA and unstructured N- terminal tails which can be highly modified (Strahl and Allis, 2000). Individual nucleosomes are packaged together by the linker histone H1 resulting in further compaction of the genetic information into chromatin. This organization of the genome is critical for protection against DNA damage and regulation of gene expression.

Histones fall into two categories, replication-dependent and –independent. The replication-dependent histones consist of the core proteins H2A, H2B, H3, H4 and the linker H1 and are encoded by non-polyadenylated mRNAs. These proteins are synthesized during S-phase of the cell cycle, in concert with DNA replication, and assembled into new chromatin structures. The replication-independent or

variant histones are encoded by polyadenylated mRNAs and synthesized throughout the cell cycle. They participate in a variety of pathways including transcription. activation (H3.3), DNA damage response (H2A.X) and kinetochore assembly (CENPA) (Sarma and Reinberg, 2005). As the genome is duplicated during S phase of the cell cycle, newly synthesized histones must be assembled into nucleosomes, a process mediated by chromatin remodeling factors and histone chaperones. The replication-coupling assembly factor (RCAF), composed of chromatin-assembly factor-1 (CAF-1) and anti-silencing factor-1 (ASF1), directs formation of nascent H3-H4 dimers into chromatin (Tyler et al., 1999). CAF-1 is a multi-subunit complex with three polypeptides (p150, p60 and p48) which interacts with proliferating cell nuclear antigen (PCNA) (Shibahara and Stillman, 1999). PCNA is the sliding clamp which forms around the replication fork and promotes processivity of DNA polymerase during DNA synthesis. The result of this interaction is a functional coupling of DNA replication and chromatin assembly. A number of chaperones have also been shown to mediate the exchange of core histones with variant histones during transcription and DNA damage repair (Eitoku et al., 2008).

### **Histone Gene Organization and Histone Locus Bodies**

Eukaryotic replication-dependent histone genes (I will refer to these replication-dependent histone genes simply as histone genes throughout) generally display a clustered organization. In yeast the genes encoding histone proteins which dimerize with one another, H2A and H2B and H3 and H4 are present in pairs that are transcribed divergently from a common promoter (Hereford et al., 1979). In *Drosophila* all five histone genes are contained in a 5 kb unit on the long arm of

chromosome 2 which is repeated ~100 times. Here again H2A-H2B and H3-H4 are divergently transcribed from opposite strands with only very short distances between their transcription start sites (Lifton et al., 1978; Pardue et al., 1977). The only *Drosophila* histone H1 gene is also present in this repeat. Mammals display a more complex organization of replication-dependent histones, and there are more than 75 genes encoding these five proteins with no repeating structure. Human histone genes are located in two clusters on chromosomes 1 and 6 (Marzluff et al., 2002).

In higher eukaryotes histone genes associate with subnuclear bodies which coordinate factors required for histone gene expression. Mammalian histone genes are localized close to Cajal Bodies (CBs) which contain the U7 snRNP (Shopland et al., 2001) and NPAT (Ma et al., 2000). CBs are sites of spliceosomal snRNP assembly (Carmo-Fonseca et al., 1992). CBs also contain coilin, which appears to function in a structural role, and Flice-associated huge protein (FLASH) which is required for processing (Barcaroli et al., 2006a; Barcaroli et al., 2006b). In *Drosophila* the histone genes associate with the Histone Locus Bodies (HLBs) which contain the U7 snRNP as well as components of the cleavage factor and are distinct structures from CBs (Liu et al., 2006). Mammalian HLBs may be defined as structures containing U7 snRNP and NPAT that are adjacent to histone genes (Ghule et al., 2008).

### **Histone Gene Expression**

Expression of histone genes is highly regulated and tightly linked to the cell cycle. In mammalian cells assembly of newly-replicated DNA into chromatin

requires an estimated  $10^8$  histones for each cell division. Regulation of histone biosynthesis is critical as production of too many histones results in chromosome loss (Meeks-Wagner and Hartwell, 1986). Histone genes are transcribed throughout the cell cycle, although their transcription rates increase 3-5 fold at the beginning of S-phase (DeLisle et al., 1983). A positive transcription factor, p220<sup>NPAT</sup> (NPAT), is present in CBs throughout the cell cycle. Phosphorylation of NPAT by cyclin E/Cdk2 as cells approach S-phase results in increased expression of histone mRNAs (Ma et al., 2000; Zhao et al., 2000). Concomitant with increased transcription is an elevated level of histone pre-mRNA processing which stabilizes the mRNAs and results in a 35-fold increase in histone mRNA levels (Fig. 1). This increase in histone pre-mRNA processing at the onset of S-phase is accomplished via upregulation of SLBP levels (Whitfield et al., 2000). Thus, throughout S-phase, histone genes are highly transcribed, their pre-mRNAs efficiently processed and packaged into mature messenger ribonucleoproteins (mRNPs) with SLBP, at which point they are ready for export to the cytoplasm and translation. At the end of S-phase, both histone mRNAs and SLBP are degraded. Histone mRNAs are degraded as a result of stopping DNA synthesis, and SLBP is degraded as a result of phosphorylation by cyclinA/Cdk1 (Koseoglu et al., 2008).

### **Replication-dependent Histone mRNAs**

Metazoan replication-dependent histone mRNAs have a 7-methyl-guanosine cap on the 5' end like polyadenylated mRNAs. However, histone mRNAs are unique in that they contain no introns and are not polyadenylated, ending instead in a conserved stemloop (SL) structure. This SL, located 30-50 nts 3' of the translation

stop codon, is highly conserved and consists of a 6 nt stem and a 4 nt loop (Fig. 2). The SL is bound by the stemloop binding protein (SLBP) which participates in all aspects of histone mRNA metabolism (see below). Variant histone mRNAs are polyadenylated like all other mRNAs.

### **Mammalian Histone pre-mRNA 3' End Formation**

Cleavage and polyadenylation of most pre-mRNAs is guided by two cis-acting elements, the polyadenylation signal (AAUAAA) 5' to the cleavage site and a G/U rich element 3' to the cleavage site. These sequences bind to the cleavage and polyadenylation specificity factor (CPSF) complex and the cleavage stimulation factor (CstF) complex, respectively, to position the cleavage factor (Mandel et al., 2008). Cleavage requires a large number of additional polypeptides although the cleavage reaction is catalyzed by CPSF73. Analogously, histone pre-mRNA processing is mediated by two cis-acting elements; the SL 5' to the cleavage site, which is bound by SLBP (Fig. 2), and a purine-rich histone downstream element, which base pairs with the U7 snRNA component of the U7 small nuclear ribonucleoprotein (U7 snRNP) (Mowry and Steitz, 1987). Cleavage of both classes of pre-mRNAs preferentially occurs after the CA dinucleotide. SLBP is a 32 kDa protein with a unique 73 aa RNA binding domain (RBD) (Wang et al., 1996). Once SLBP binds to the SL, the U7 snRNP is recruited to the HDE and its interactions with the HDE are stabilized by SLBP (Dominski et al., 1999). U7 snRNP is composed of U7 snRNA and a heteroheptameric ring of proteins. U7 snRNP contains SmB, SmD3, SmE, SmF, and SmG, like spliceosomal snRNPs, but SmD1 and SmD2 are replaced by the Sm-like proteins Lsm10 and Lsm11 (Pillai et al., 2001; Pillai et al.,

2003). The U7 snRNA itself is similar to spliceosomal snRNAs with a 5' 2,2,7 trimethylguanosine cap and a stemloop at the 3' end. In mammals, ZFP100 acts as a bridging factor, binding both SLBP and Lsm11 to stabilize U7 snRNP binding and help recruit the cleavage factor (Dominski et al., 2002; Wagner et al., 2006). ZFP100 is the limiting factor in mammalian histone pre-mRNA processing and is required for entry into S-phase (Wagner and Marzluff, 2006). The cleavage reaction is carried out by a subset of polyadenylation factors consisting minimally of CPSF73, CPSF100, and Symplekin (Fig. 3, top) (Wagner et al., 2007). CPSF73 has been identified as the endonuclease for both polyadenylated (Mandel et al., 2006) and histone mRNAs (Dominski et al., 2005b). CPSF73 and CPSF100 are members of the metallo- $\beta$ -lactamase family of zinc-dependent endonucleases and are both required for cleavage, though CPSF100 lacks residues critical for catalysis (Kolev et al., 2008). After cleavage, SLBP remains associated with the mRNA and escorts it into the cytoplasm (Fig. 4).

### **Histone pre-mRNA Processing in *Drosophila***

Histone pre-mRNA processing is remarkably well conserved from *Drosophila* to mammals (Fig. 3). Cleavage is mediated by SLBP and the U7 snRNP in both, although it does not require SLBP in mammals, while in *Drosophila* SLBP is required in vitro (Dominski et al., 2005a). No orthologue of ZFP100 has been identified in *Drosophila* (Wagner et al., 2007). Each of the canonical histone genes contains multiple poly(A) sequences downstream of the normal cleavage sites which are used when histone pre-mRNA processing factors are absent, such as in SLBP (Lanzotti et al., 2002) and U7 snRNA (Godfrey et al., 2006) null mutant flies. Interestingly, each



of these mutants displays a different developmental phenotype, with the SLBP null being lethal and the U7 snRNA mutant being sterile, likely as a result of a later onset of histone misprocessing in the U7 snRNA mutant. A characteristic of *Drosophila* histone genes is the presence of strong RNA polymerase pause sites just downstream of the cleavage site in *Drosophila* (Adamson and Price, 2003) which may allow for assembly of the processing machinery on the nascent RNA.

### **Translation and Degradation of Histone mRNAs**

Polyadenylated mRNAs are bound to the cap binding complex (CBC) on their 5' end and poly(A) binding protein on their poly(A) tails; these proteins interact with eIF4G, resulting in a circular RNA which is ready for multiple rounds of efficient translation. Histone mRNAs are translated by a similar mechanism; their 5' ends are bound by the CBC, however, since they lack poly(A) tails, a different mechanism circularizes the mRNA. SLBP is associated with the 3' end of the mRNA after the cleavage reaction and binds the SLBP interacting protein 1 (SLIP1) (Cakmakci et al., 2008) which binds to eIF4G resulting in a circular mRNA which can be efficiently translated (Fig. 4).

At the end of S-phase or when DNA synthesis is inhibited, histone mRNAs are rapidly degraded (Fig. 4) in a manner that is dependent upon translation; treatment with cycloheximide, a general translational inhibitor, stabilizes the mRNAs (Kaygun and Marzluff, 2005b). This mechanism of degradation also requires the SL and SLBP (Kaygun and Marzluff, 2005b; Pandey and Marzluff, 1987) and requires that the stemloop to be close to the termination codon. SLBP recruits Upf1 and a

terminal uridylyl transferase (TUTase) to the 3' ends of histone pre-mRNAs which add an 8-12 nt oligo(U) tail. The oligo(U) likely serves as a platform for binding of the Lsm1-7 complex (Mullen and Marzluff, 2008). This complex, in turn, recruits the decapping enzymes freeing the 5' end for degradation by the 5'-3' exonuclease Xrn1. There is also a role for the exosome in 3'-5' degradation of histone mRNAs, which results in a bidirectional decay mechanism for histone mRNAs in mammals.

### **Cleavage and Polyadenylation of pre-mRNAs**

Transcription of protein coding genes in eukaryotes yields pre-mRNAs which must be processed into mature transcripts, ready for export and capable of being translated. Processing events include 5' capping, splicing (in many cases) and proper formation of the 3' end. For all pre-mRNAs except histone mRNAs, proper formation of the 3' end requires endonucleolytic cleavage and addition of a long (~250nt) poly(A) tail. Polyadenylation is required for nuclear export (Huang and Carmichael, 1996), stability (Coller et al., 1998) and efficient translation (Preiss and Hentze, 1998) of mRNAs. Cleavage and polyadenylation requires a large protein complex with as many as 85 polypeptides (but no RNAs) playing roles (Shi et al., 2009).

Cleavage and polyadenylation is guided by multiple cis acting elements most notably the poly(A) site (PAS), AAUAAA, and the downstream sequence element (DSE), which is G/U rich. These sites interact with the highly conserved CPSF and CstF complexes respectively. The CPSF complex contains five subunits in mammals, CPSF30, CPSF73, CPSF100, CPSF160, and Fip1p (Fig. 5). CPSF160,

the largest subunit of the CPSF complex, interacts directly with the AAUAAA to facilitate cleavage (Zhao et al., 1999a), and also with CPSF100 (Murthy and Manley, 1995) and Fip (Murthy and Manley, 1992). CPSF73 and CPSF100 are members of the metallo- $\beta$ -lactamase family of endonucleases. CPSF73 is the nuclease for cleavage of polyadenylated mRNAs, and CPSF100 is also required for cleavage, though its nuclease domain is inactive (Kolev et al., 2008). CPSF30 is required for cleavage and polyadenylation and binds a U-rich sequence element upstream of the PAS (Barabino et al., 2000). Fip1p, factor interacting with Pap1p, was initially identified as a protein which interacts with the poly(A) polymerase, Pap1p (Preker et al., 1995). Fip1p also interacts with CPSF160 (Kaufmann et al., 2004), CPSF30, and CstF77 (Helmling et al., 2001), directing Pap1p to the correct poly(A) site. Another component of the CPSF complex required for proper 3' end formation and polyadenylation is Symplekin. Pta1p, the yeast orthologue of Symplekin was originally identified as required for pre-tRNA processing (O'Connor and Peebles, 1992a), while Symplekin itself was found to be associated with tight junctions (Keon et al., 1996). Symplekin may be a scaffolding protein (Fig. 5) as it binds to a large number of proteins including CPSF73 (Zhelkovsky et al., 2006), CPSF100 (Kyburz et al., 2003), CstF64 (Takagaki and Manley, 2000a) and the RNA Pol II CTD phosphatase Ssu72 (Skaar and Greenleaf, 2002). Symplekin is also the heat labile factor required for histone pre-mRNA processing (Kolev and Steitz, 2005).

The CstF complex is composed of 3 subunits, CstF50, CstF64, and CstF77 (Fig. 5). The CstF64 subunit binds directly to the G/U rich DSE (Takagaki and Manley, 1997) and also to CstF77 (Hatton et al., 2000) and. CstF77, identified as

the suppressor of forked (*su(f)*) in *Drosophila*, assists in poly(A) site selection as mutations in this gene result in poly(A) site switching (Juge et al., 2000). CstF77 exists as a homodimer, and this dimerization appears to be critical for its function (Bai et al., 2007; Benoit et al., 2002). Additionally, CstF77 binds CstF50 and CPSF160, linking the complexes upstream and downstream of the cleavage site. The final component of the CstF complex, CstF50, binds CstF77 as mentioned and also homodimerizes. The entire CstF complex is likely present as a homodimer in polyadenylation (Mandel et al., 2008).

Higher eukaryotes require additional factors termed mammalian cleavage factor (CF) I<sub>m</sub> and II<sub>m</sub> which further aid in cleavage and polyadenylation as well as linking this process with others such as splicing and transcription. CFI<sub>m</sub> is composed of three subunits of 25, 59 and 68 kDa (Fig. 5), and it functions as a heterodimer of the 25 kDa subunit and one of the others, with different combinations contributing to various activities. The 25 and 68 kDa subunits reconstitute *in vitro* cleavage activity in the presence of partially purified 3' processing complexes (Ruegsegger et al., 1998). All three subunits crosslink to RNA in a sequence dependent manner (Ruegsegger et al., 1998) with a preference for UGUAA (Brown and Gilman, 2003), a sequence generally found upstream of poly(A) sites. CFI<sub>m</sub> aids in recognition of both canonical and non-canonical poly(A) sites (Venkataraman et al., 2005). CFII<sub>m</sub> is composed of Pcf11 and Clp1 both of which are highly conserved among all eukaryotes (Fig. 5). Mutations in yeast Pcf11 result in cell death without affecting 3' end formation (Sadowski et al., 2003), but instead result in aberrant transcription termination (Zhang et al., 2005). Clp1 interacts with Pcf11 and is

required for cleavage but not polyadenylation (Kyburz et al., 2003). Clp is an RNA 5' kinase also important for tRNA splicing and activation of siRNAs (Weitzer and Martinez, 2007).

Eukaryotic pre-mRNA 3' end formation is an incredibly complex mechanism involving many proteins, some of which have, undoubtedly, yet to be identified, and yet serve to cleave a single phosphodiester bond. This process is critical for cell viability as improperly processed mRNAs are not exported from the nucleus efficiently or translated. Recent work has demonstrated that this process and others are coordinated with transcription to ensure fidelity.

### **Cotranscriptional Processing of pre-mRNAs**

Transcription and RNA processing were traditionally thought of as discrete processes, that is to say, only after a full length transcript was produced, was it released from the template then cleaved and polyadenylated and then spliced. There has been mounting evidence over the last 20 years, however, that these phenomena are intertwined. The current model of RNA processing is a cotranscriptional one in which nascent RNA molecules are processed as they are extruded from transcribing RNA Pol II.

Targeting of pre-mRNA processing factors to nascent transcripts occurs in one of two ways; 1) direct recruitment via interaction with RNA pol II (usually with the C-terminal domain or CTD) or 2) association of subnuclear structures such as CBs or splicing speckles with sites of transcription or possibly a combination of the two. The CTD of RNA Pol II is disordered and composed of a heptad with a consensus

sequence of YSPTSPS which is repeated 26 times in yeast and 52 times in mammals. Phosphorylation can occur on each of the serines (S2, S5, and S7) and each has a different role in recruitment of processing factors to RNA Pol II. Additionally, a peptidyl-prolyl isomerase, Pin1, binds to and acts upon the phosphorylated CTD, greatly expanding the number of possible conformations this structure can adopt.

Phosphorylation of the CTD of RNA Pol II occurs in a distinct pattern throughout the bodies of genes (Fig. 6). Prior to initiation, the CTD is not phosphorylated. During initiation of transcription at the 5' ends of genes the CTD is primarily phosphorylated on S5 by the Cdk7 subunit of the general transcription factor TFIIF (Cismowski et al., 1995; Komarnitsky et al., 2000). At the transition to the elongation step of transcription, or promoter clearance, phosphorylation of the S2 position by positive elongation factor b or P-TEFb (cyclin T/Cdk9) occurs (Marshall et al., 1996). This phosphorylation occurs throughout the coding region of genes (Komarnitsky et al., 2000) and in vitro is enhanced on CTD repeats already phosphorylated on S5 (Jones et al., 2004), fittingly nicely with the model of phosphorylation state switching during active transcription. Phosphorylation of S7 has been claimed to occur on a class of genes which do not encode mRNAs, but rather snRNAs which have a unique 3' end. Phosphorylation may be required for proper 3' end formation as it recruits the integrator complex which contains the cleavage factor, however, the kinase responsible for it remains unknown (Egloff et al., 2007a).

Examples of interactions between RNA processing factors and the CTD of RNA Pol II abound (Neugebauer, 2002; Phatnani and Greenleaf, 2006) so in this section I will limit my discussion to a few prominent examples for which there is compelling evidence *in vivo*. Recruitment of the 5' RNA capping enzyme (HCE) to nascent RNAs is mediated by S2 phosphorylation of the CTD. Interestingly, although the enzyme can bind either the S2 or S5 form of the CTD *in vitro*, the S2 CTD actually stimulates its guanylyl transferase activity (Ho and Shuman, 1999). The evidence for cotranscriptional splicing is extensive (Kornblihtt et al., 2004), however, there is little evidence for direct interactions between spliceosomal proteins or snRNPs and the CTD, with the lone example being the yeast protein Prp40p which is a component of the U1 snRNP (Morris and Greenleaf, 2000). While direct evidence for binding is lacking, indirect evidence is substantial (Misteli and Spector, 1999; Moore et al., 2006). A large class of splicing factors which are not catalytic and contain serine-arginine rich domains, the SR proteins, have been shown to interact with the CTD in yeast two hybrid screens and by coimmunoprecipitation (Yuryev et al., 1996). Cotranscriptional splicing is clearly vital to the cell as changes in transcription kinetics alter exon inclusion rates (Listerman et al., 2006) and loss of the splicing factor ASF/SF2, which is recruited cotranscriptionally, results in genomic instability due to the formation of R-loops (highly unstable DNA-RNA hybrids) (Li and Manley, 2005).

As previously discussed, proper 3' end processing of pre-mRNA is crucial for proper gene regulation and this process also occurs cotranscriptionally. Both the CPSF and CstF complexes contain components which interact directly with the S2

phosphorylated CTD of RNA Pol II. Initial insights into coupling of 3' end formation and transcription revealed that not only was the CTD required for proper cleavage and polyadenylation, but that several factors (CstF50 and CstF77) bind directly to the phosphorylated CTD resulting in the CstF complex binding to phosphorylated CTD columns (McCracken et al., 1997). The yeast orthologue of CPSF160 has also been shown to bind the CTD directly (Dichtl et al., 2002), providing evidence for direct recruitment of both major cleavage complexes to RNA via the polymerase. Additionally, transcription termination factors (see below) such as Xrn2 (Kaneko et al., 2007) and Pcf11 (Sadowski et al., 2003; Zhang and Gilmour, 2006) are recruited cotranscriptionally demonstrating that RNA Pol II is only capable of terminating after forming the proper 3' end.

### **Transcription Termination**

One area of cotranscriptional regulation of gene expression which lags behind some others is our understanding of the mechanism of transcription termination. After initiation of transcription, capping of the mRNA, conversion of RNA Pol II to a processive form for elongation by PTEF-b, identification of exons and 3' end formation, the polymerase is still engaged on the DNA and transcribing. Two models for transcription termination in eukaryotes (different from bacterial rho termination) have emerged recently termed the allosteric and torpedo models and ample evidence exists for both as well as a synergistic mechanism combining the two. Both rely on the observation that interfering with polyadenylation prevents transcription termination. In the allosteric model, after cleavage and polyadenylation, Pcf11 bridges the RNA to the CTD of RNA Pol II releasing the



transcript from the template and resulting in a conformational change in RNA Pol II which reduces its processivity and eventually leads to termination (Fig. 7A). Pcf11 binds to RNA and the CTD through same domain, and competition for this site may facilitate dismantling of the transcription complex (Hollingworth et al., 2006; Zhang et al., 2005). The torpedo model posits that the free 5' end of the nascent RNA generated by 3' end formation is a substrate for the 5'-3' exonuclease Xrn2 which degrades this RNA until it "catches up" with the polymerase and aids in dissociating it from the template (Fig. 7B). Xrn2 localizes to the 3' ends of genes in vivo by ChIP and interacts with 3' processing factors (Kaneko et al., 2007), providing a potential mechanism for recruitment. Recent evidence demonstrates that recruitment of Pcf11 or Rat1 (the yeast orthologue of Xrn2) enhances recruitment of the other factor. This observation supports a hybrid of the two models (Luo et al., 2006a) where both Pcf11 interacting with the CTD and degradation of the 3' cleavage product contribute to termination of transcription, as does pausing of RNA Pol II.

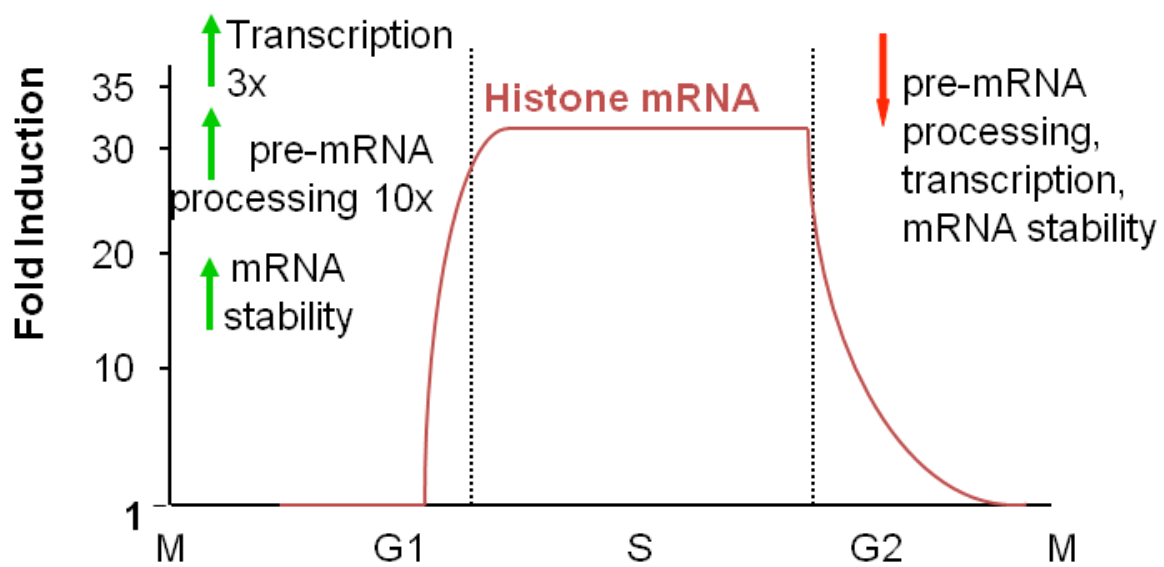
### **Summary**

Proper regulation of histone biosynthesis is critical for proper cell division and gene regulation. These extremely abundant proteins are only made during S-phase and regulation of histone protein synthesis is mediated primarily at the RNA level. Modest increases in transcription contribute to the increase in histone mRNA present in S-phase in mammalian cells, but the major contributor is activation of 3' end processing. The 3' end processing reaction is a single endonucleolytic cleavage downstream of the highly conserved stemloop and occurs in a fashion analogous to polyadenylated mRNAs with cis elements upstream and downstream of the

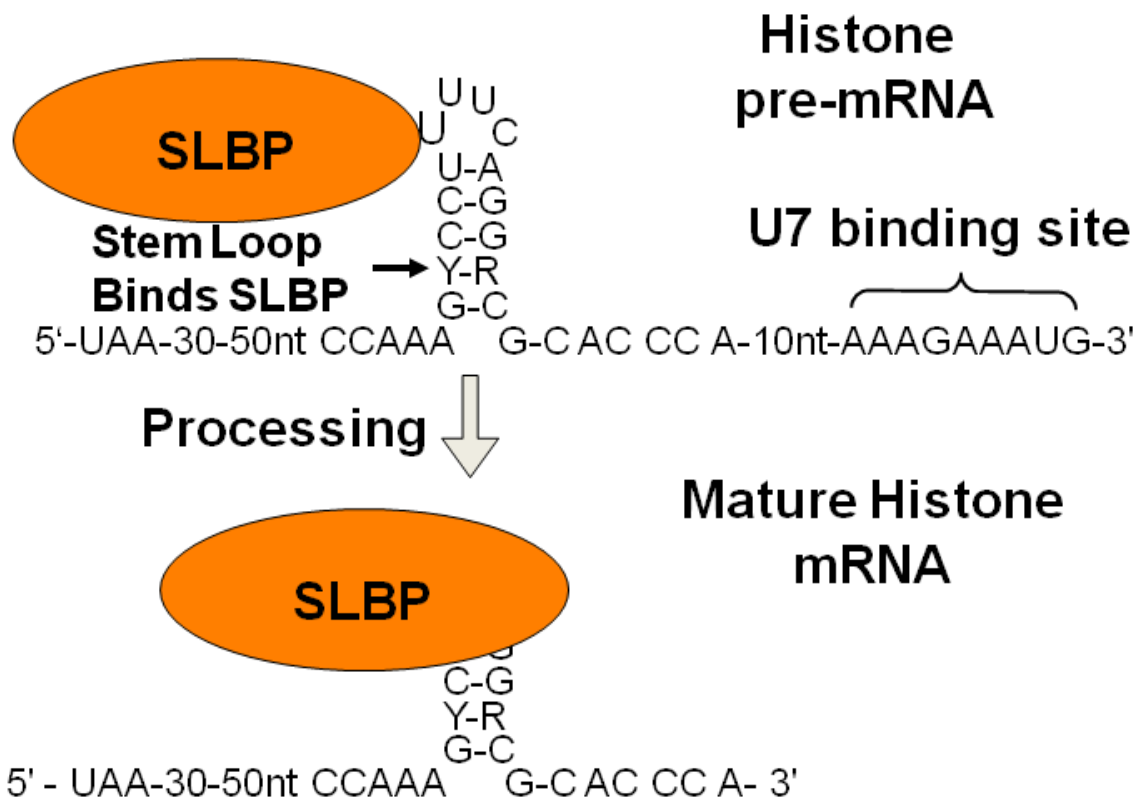
cleavage site positioning a cleavage factor. Recent studies have shown that histone pre-mRNA processing is not as different from cleavage and polyadenylation as we once believed. Cleavage of histone pre-mRNA actually uses some of the same proteins as polyadenylated mRNA, including the endonuclease CPSF73. What is not known is exactly which of these proteins are required for processing of endogenous histone mRNAs in vivo. We also do not know how all of these proteins interact with one another and more importantly how they interact with histone specific processing factors to generate specificity. Another important question concerning histone pre-mRNA processing is whether this occurs co- or post-transcriptionally as there are currently conflicting reports. Furthermore, the mammalian 5'-3' exonuclease Xrn2 is recruited to downstream cleavage products by the CstF complex to promote transcription termination on genes encoding polyadenylated mRNAs. As we have no evidence of a role for CstF in histone pre-mRNA processing, this raises the question of how transcription of histone genes is terminated.

Chapter 2 of this dissertation examines the role of the *Drosophila* poly(A) processing apparatus in histone pre-mRNA processing and presents evidence that this reaction occurs cotranscriptionally. Chapter 3 presents work pertaining to the mechanism of transcription termination on mammalian histone genes, specifically showing that Xrn2 is not a major player in the pathway. In Chapter 4, I describe an unexpected function for SLBP in histone mRNA export. And the final chapter (Chapter 5) will summarize these data and place them into the context of the field of mRNA processing.

**Figure 1. Histone mRNA and the cell cycle.** Histone mRNA biosynthesis peaks in S-phase, the result of synergistic increases in transcription, 3' end formation and stability. At the end of S-phase, these same processes are down-regulated and the mRNA is actively degraded. Together these mechanisms restrict histone protein synthesis to S-phase.

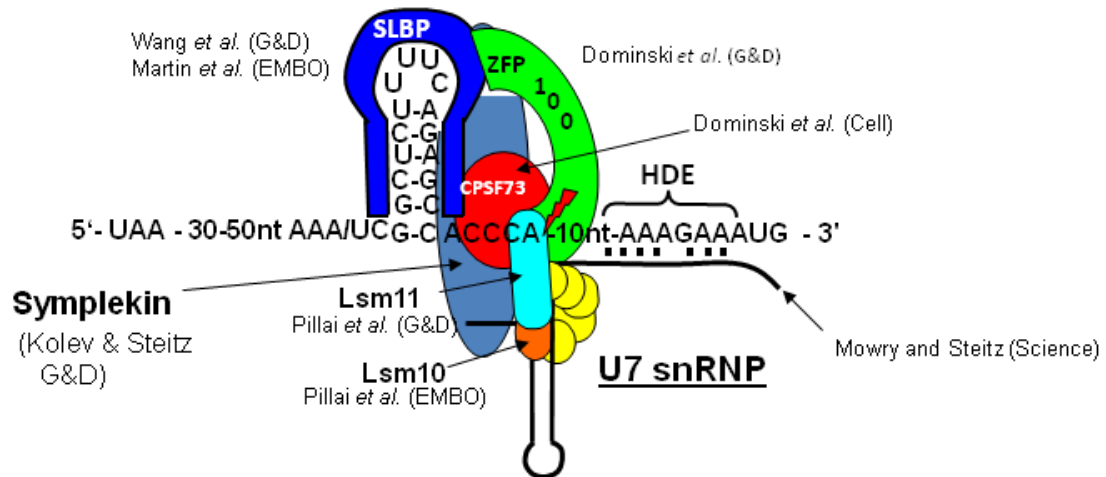


**Figure 2. Schematic of histone pre-mRNA processing.** Histone genes contain no introns, therefore, the only processing event in the maturation process is a single endonucleolytic cleavage. This cleavage is mediated by two cis elements the SL, upstream of the cleavage site, and the HDE, downstream of the cleavage site. SLBP binds the SL and U7 snRNA base pairs with the HDE.

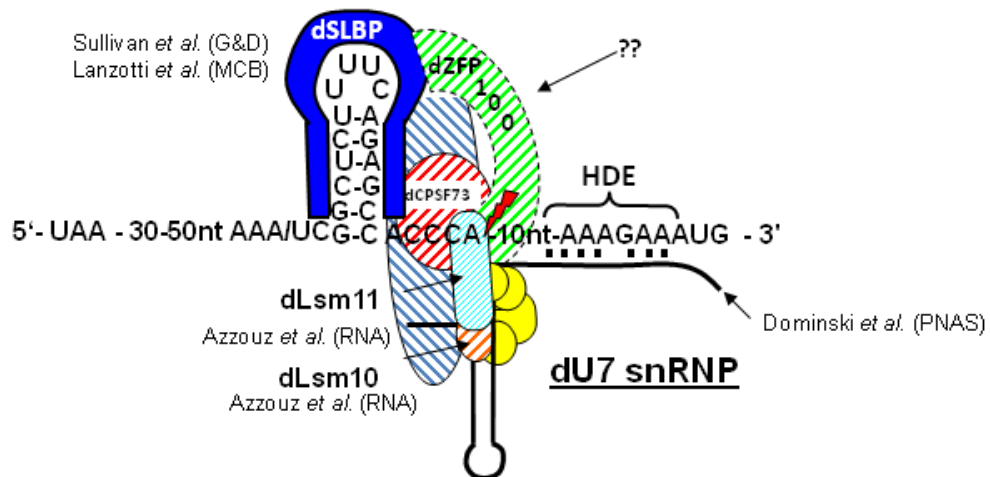


**Figure 3. Conservation of histone pre-mRNA processing machinery between mammals and flies.** A schematic showing known histone pre-mRNA processing components in *Homo Sapiens* and *Drosophila Melanogaster*.

## Homo Sapiens



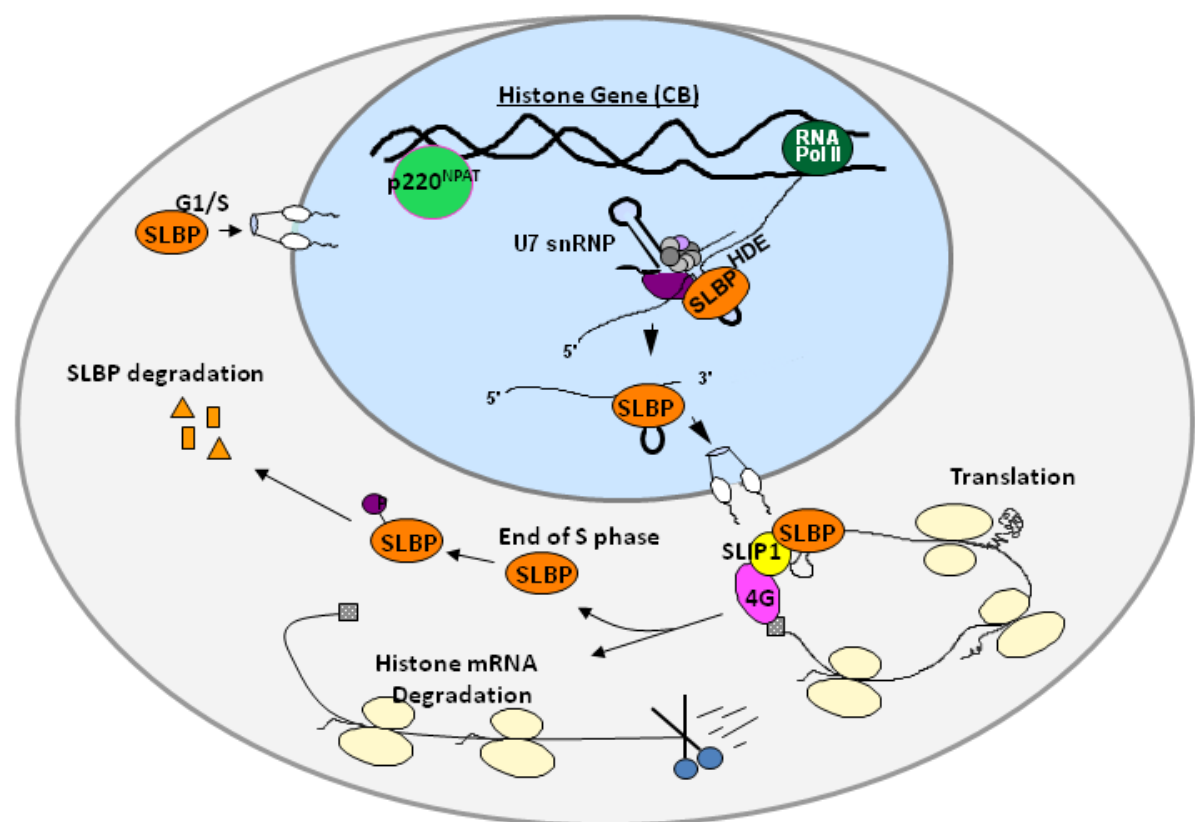
## Drosophila Melanogaster





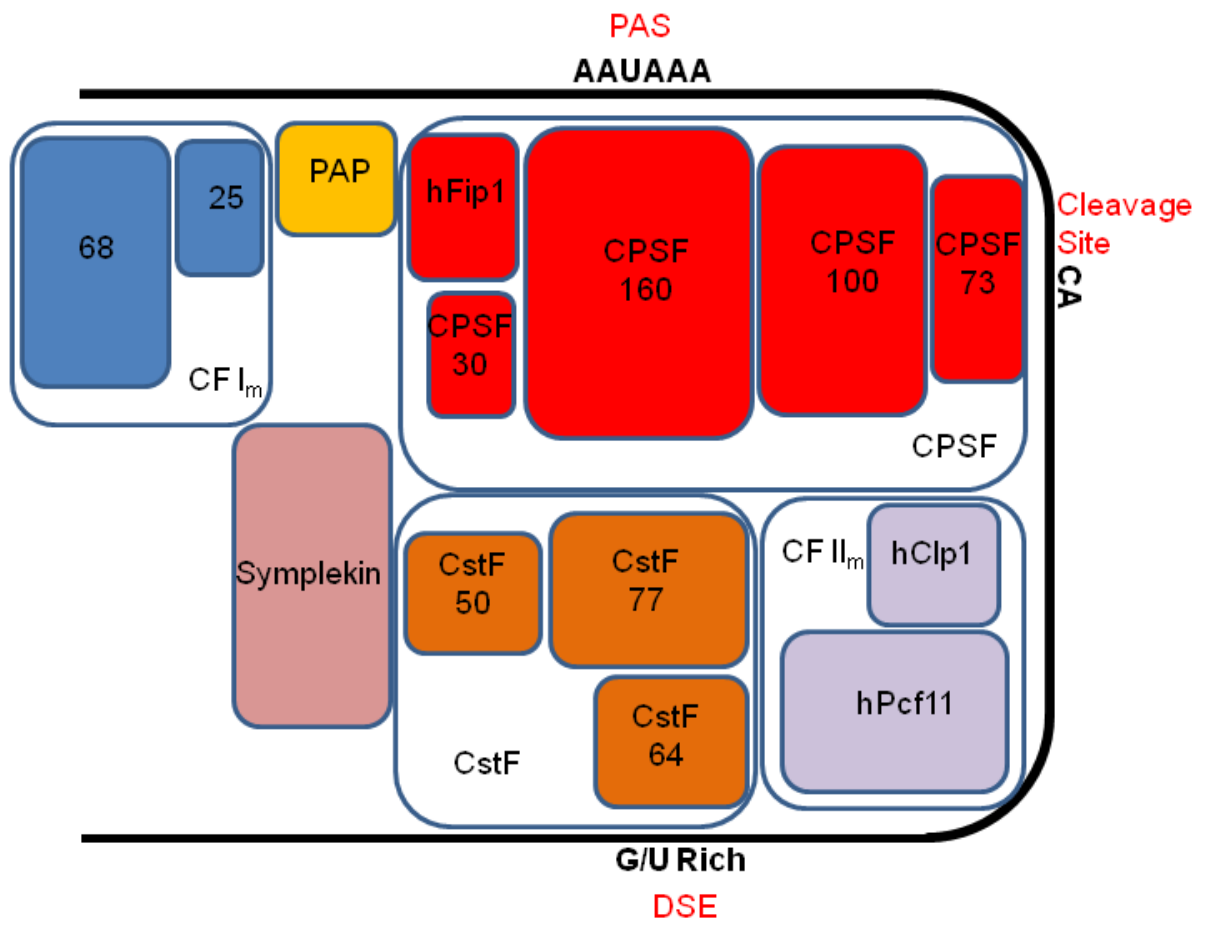
**Figure 4. Global view of histone mRNA metabolism in mammalian cells.**

Histone genes are transcribed at or near Cajal Bodies. Processing of the pre-mRNAs is accomplished by SLBP, the U7 snRNP, and the cleavage complex. SLBP remains associated with the mature mRNA and escorts it into the cytoplasm where it is active in translation. At the end of S-phase, both histone mRNAs and SLBP are degraded.

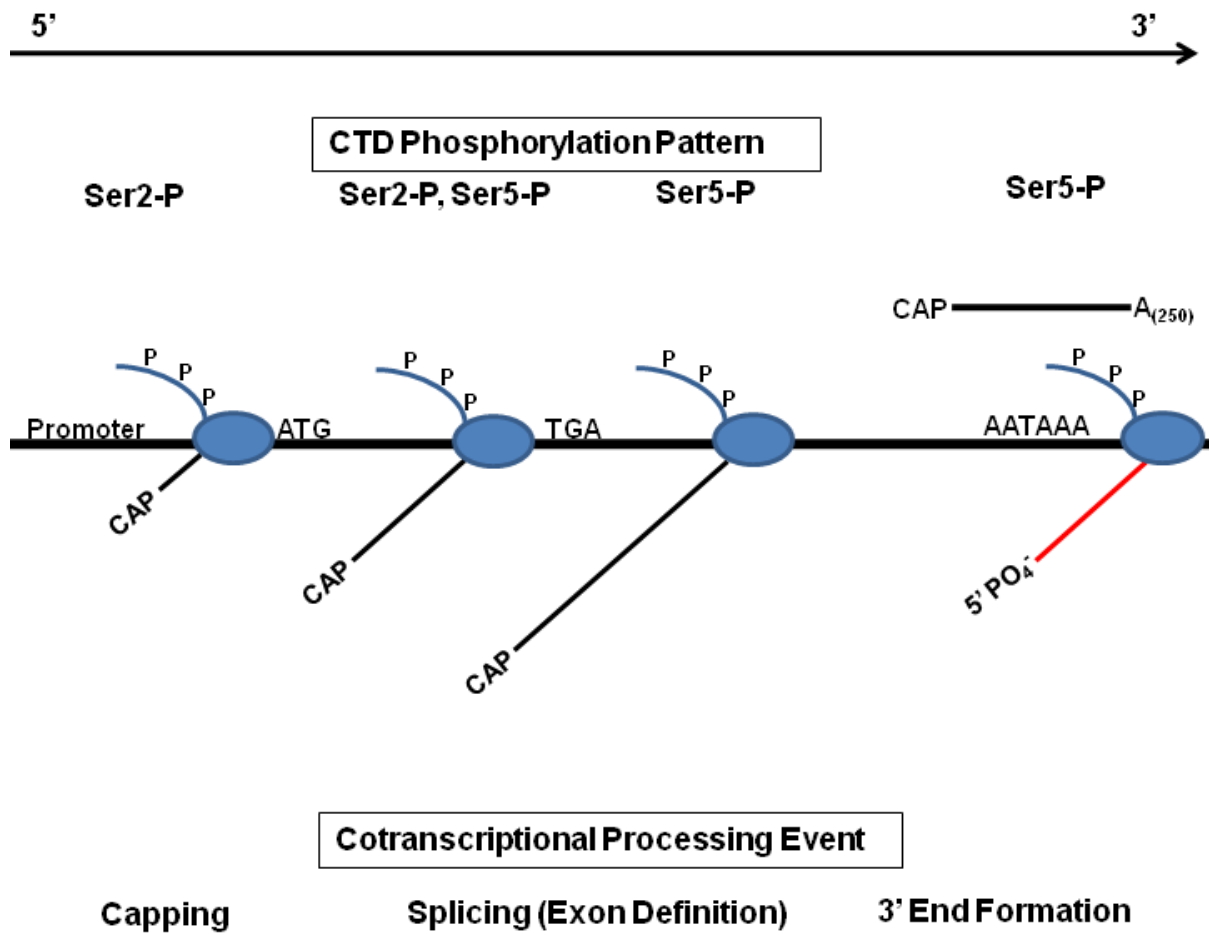


**Figure 5. Schematic of mammalian cleavage and polyadenylation machinery.**

Adapted from (Mandel et al., 2008). Schematic of various cleavage and polyadenylation complexes and the DNA sequences that they bind.

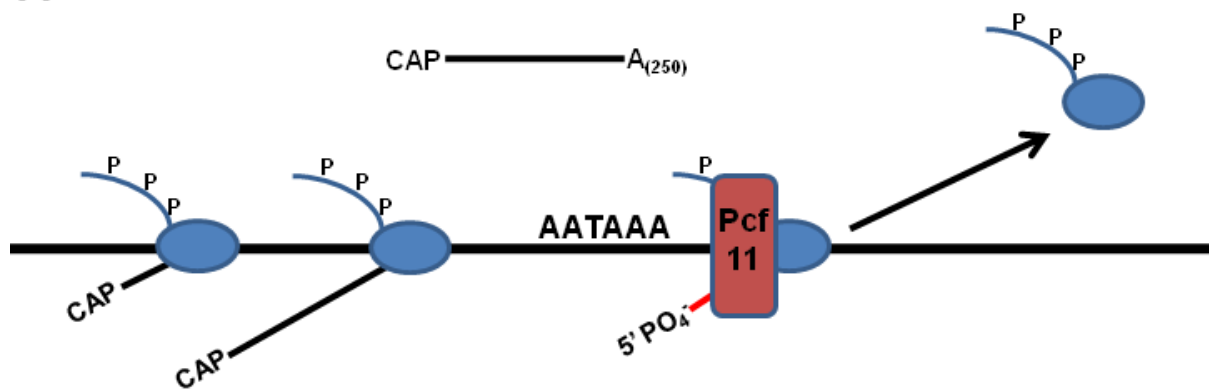


**Figure 6. Cotranscriptional pre-mRNA processing.** Many RNA processing events take place while transcription is occurring rather than after transcription is completed and the factors responsible for these reactions are recruited via the CTD of RNA Pol II. Differential phosphorylation of the CTD of RNA Pol II on Ser2 and Ser5 recruit different proteins at various times throughout transcriptions. As the polymerase proceeds along the gene and the phosphorylation state changes, mRNAs are capped, spliced, cleaved and polyadenylated.

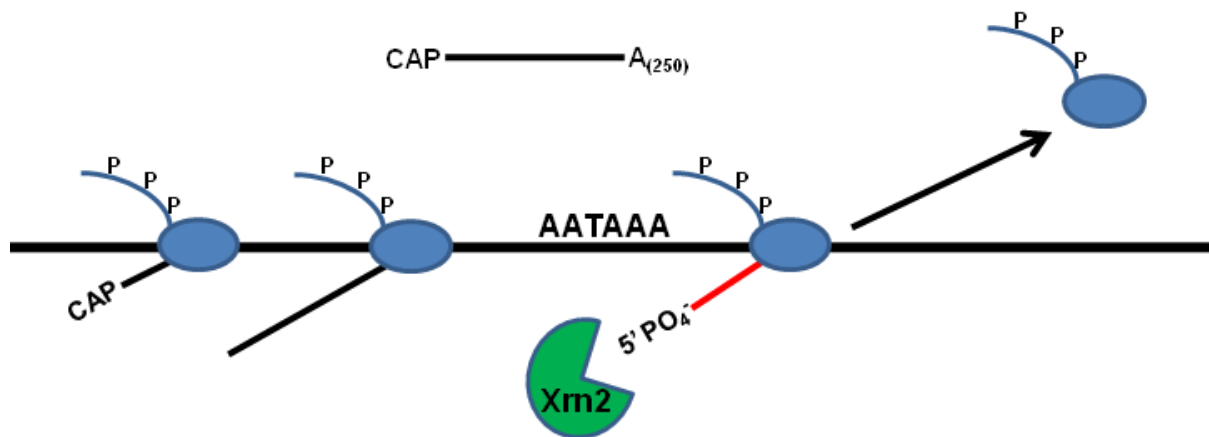


**Figure 7. Transcriptional termination models.** Schematics of two current models of transcription termination. **(A)** Allosteric model. After cleavage and polyadenylation, the termination factor Pcf11 bridges the CTD of RNA Pol II to the nascent RNA causing a conformational change and dismantling the transcription complex. **(B)** Torpedo model. The free 5' phosphate generated by the 3' end formation reaction serves as a substrate for a nuclear 5'→3' exonuclease (Xrn2 in mammals or Rat1 in yeast) which degrades the RNA until it catches up to the polymerase and knocks it off of the DNA template.

**A**



**B**





## **Chapter 2**

### **A Subset of Polyadenylation Factors is Required for Cotranscriptional Histone pre-mRNA Processing in *Drosophila***

#### **Introduction**

Metazoan histone mRNAs are unique among eukaryotic mRNAs in that they contain no introns and end in a conserved stem-loop structure rather than a poly(A) tail. Formation of a mature histone mRNA thus requires only a single endonucleolytic cleavage. This cleavage is directed by two cis elements, a highly conserved stem loop (SL) 5' of the cleavage site and a less conserved purine-rich histone downstream element (HDE) 3' of the cleavage site. There are two histone specific RNA processing factors, the stem loop binding protein (SLBP) which binds the SL and the U7 RNA of the U7 small nuclear ribonucleoprotein (U7 snRNP) which base pairs with the HDE (Marzluff et al., 2008). The U7 snRNP consists of the U7 snRNA and a heteroheptameric ring of Sm proteins in which the spliceosomal Sm proteins Smd1 and Smd2 are replaced by Lsm10 and Lsm11 (Pillai et al., 2003). Cleavage is catalyzed by CPSF73 (Dominski et al., 2005b) and Symplekin has been

implicated as the scaffold which coordinates formation of the cleavage complex (Kolev and Steitz, 2005). Following processing, the mature mRNA is escorted into the cytoplasm by SLBP (Sullivan et al., 2009) where SLBP participates in efficient translation of histone mRNA (Sanchez and Marzluff, 2002).

Cleavage and polyadenylation of all other metazoan mRNAs requires two multi-protein complexes termed the cleavage and polyadenylation specificity factor (CPSF) and the cleavage stimulation factor (CstF), which recognize signals upstream and downstream of the cleavage site respectively. CPSF is composed of CPSF30, CPSF73, CPSF100 and CPSF160. These proteins have been shown to interact with one another [reviewed in (Mandel et al., 2008)] and with the AAUAAA polyadenylation signal which is recognized by CPSF160 (Keller et al., 1991; Murthy and Manley, 1995). Both CPSF73 and CPSF100 have putative  $\beta$ -lactamase domains, and CPSF73 has been described as the endonuclease for both poly(A) (Mandel et al., 2006; Ryan et al., 2004) and histone mRNAs (Dominski et al., 2005b; Yang et al., 2009). CPSF100 has also been shown to play an important role in these reactions (Kolev et al., 2008) although it lacks critical residues required for catalysis. CstF is comprised of three polypeptides CstF50, CstF64, and CstF77, with CstF64 binding the downstream GU-rich element required for polyadenylation (Yoshio and Manley, 1997).

Symplekin was originally identified as a tight junction protein in mammalian cells (Keon et al., 1996) and its yeast homolog, Pta1p, was previously characterized as being essential for pre-tRNA processing (O'Connor and Peebles, 1992b). Symplekin has subsequently been shown to be involved in cleavage and

polyadenylation via interactions with both the CPSF and CstF in yeast (Preker et al., 1997; Zhao et al., 1999b) and mammals (Takagaki and Manley, 2000a; Vethantham et al., 2007). Additionally, Symplekin was defined as the elusive heat labile factor (Gick et al., 1987) required for histone pre-mRNA processing (Kolev and Steitz, 2005).

In *Drosophila*, there is only one histone gene for each of the five canonical histone proteins. These genes exist as a ~5kb unit which is tandemly repeated ~100 times on the long arm of chromosome 2. This histone gene cluster localizes to nuclear structures termed the histone locus bodies (HLBs) which are located in close proximity to Cajal bodies (CBs) (Liu et al., 2006). HLBs also contain U7 snRNP (Liu et al., 2006), Symplekin [(Wagner et al., 2007); D. Tatomer, K.D.S, M.S. and W.F.M., unpublished] and a cyclin E substrate, Mpm2 (White et al., 2007), consistent with a role for these nuclear bodies in histone pre-mRNA processing.

Various RNA processing events have been shown to occur cotranscriptionally, often as a result of the recruitment of RNA processing factors by the C-terminal domain (CTD) of RNA Pol II [reviewed in (Bentley, 2005; Buratowski, 2005a; Neugebauer, 2002; Proudfoot, 2004)]. Components of the poly(A) apparatus have been shown to bind to the CTD of RNA Pol II (Kyburz et al., 2003; McCracken et al., 1997; Phatnani and Greenleaf, 2004) suggesting that at least some components of the polyadenylation machinery are loaded onto the polymerase. In addition the CTD is important for efficient cleavage/polyadenylation reaction in vivo (Hirose and Manley, 1998). Polyadenylation factors are associated with genes

encoding mRNAs as detected by chromatin immunoprecipitation (ChIP), consistent with cleavage/polyadenylation occurring cotranscriptionally.

Histone mRNA expression is tightly regulated and coordinated with the cell cycle. At the beginning of S phase, histone mRNA levels increase ~35-fold to produce enough histones to package newly synthesized DNA. This increase in mRNA is accomplished by upregulating transcription and 3' end processing, which is absolutely dependent on both the U7 snRNP and SLBP in *Drosophila* (Dominski et al., 2005a). In vitro data strongly suggest that RNA Pol II pauses just 3' of the processing site, possibly to allow cotranscriptional assembly of the processing complex (Adamson and Price, 2003). The 3' ends of four of the histone genes are less than 500 nts from the 3' end of an adjacent gene (transcribed from the opposite strand, Fig. 1A). Thus, to prevent read-through into the adjacent gene, it is essential to efficiently terminate transcription. There are cryptic polyadenylation signals downstream of each *Drosophila* histone gene. If the processing efficiency of histone mRNAs is reduced either by mutation or knockdown of factors required for histone mRNA processing, then RNA Pol II reads through and the mRNAs become polyadenylated (Godfrey et al., 2006; Sullivan et al., 2001).

A recent RNA interference screen implicated a subset of polyadenylation factors, Symplekin, CPSF73 and CPSF100 in histone pre-mRNA processing, while other polyadenylation factors did not score in the screen (Wagner et al., 2007). To further investigate the role of these proteins in histone pre-mRNA processing, we first examined the effect of RNAi-depletion of these factors on the 3' end of endogenous histone mRNA. We carried out co-immunoprecipitation (coIP)

experiments and chromatin immunoprecipitation (ChIP), to demonstrate that Symplekin, CPSF73 and CPSF-100 are part of a core cleavage factor involved in cotranscriptional histone mRNA 3' end processing.

## MATERIALS AND METHODS

**RNAi and NE preparation.** Approximately 500 bp dsRNAs were produced by T7 transcription, DNA template digestion and annealing. RNAi was performed by adding 100 µg of dsRNA to Dmel-2 cells ( $2 \times 10^7$ ) on day 1, 200 µg on day 2 and 400 µg on day 3. Following a two-day recovery, NEs were prepared as described (Wagner et al., 2007).

**Antibodies.** A fragment of recombinant *D. Melanogaster* Symplekin (amino acids 1-767) and full length CPSF73 were expressed in E.coli and used to inoculate both rabbits and guinea pigs for antibody production (Pacific Immunology Corp.).  $\alpha$ -CPSF73 was antigen purified against full length CPSF73. A fragment of recombinant *D. Melanogaster* SLBP (amino acids 1-175) or Lsm11 (amino acids 1-100) was used to inoculate rabbits for antibody production (Proteintech Group, Inc.). The *D. Melanogaster*  $\alpha$ -CstF50 antibody (Ni et al., 2008) was a generous gift from John Lis (Cornell Univ.) and the  $\alpha$ -CstF77 antibody was from Bethyl Labs. The CPSF-100 antibody was raised against a conserved region of human CPSF-100 (Z. Dominski and X. Yang, unpublished) and cross-reacted with the *D. Melanogaster* CPSF-100 protein.

**Analysis of RNA.** The S1 nuclease assay was performed as described (Lanzotti et al., 2002) . Poly(A) mRNA was separated from total cellular RNA using the Poly(A)Purist mRNA isolation kit (Ambion).

**Immunoprecipitations(IPs).** 500 µg of NE was incubated with pre-conjugated protein-A sepharose beads in RIPA (50 mM Tris, pH 8.0, 1% NP-40, 0.25% deoxycholate, 150 mM NaCl) or a low detergent buffer (50 mM Tris, pH 8.0, 0.1% NP40, 150 mM NaCl) overnight. Washes were performed in IP buffer and the precipitated proteins were eluted into SDS loading buffer.

**Chromatin Immunoprecipitation.** ChIP was performed as described for *Drosophila* (Gilchrist et al., 2008) or human (Upstate protocol) and 15 µl of CPSF73, HA, Rpb3, hSLBP, Symplekin and CstF50 antibodies were used in each IP. qPCR was performed and quantified using SYBR green (Applied Biosystems) and an Applied Biosystems 7900HT real-time PCR machine.

Oligonucleotide sets used:

*Drosophila*

H2A F-AATGGAATATCTGGCCGCTGAGGT, R-  
CCAGTTGCAGATGACGCGGAATAA;

H2B F-CGGGCCCACATTACGATAAGTGAA, R-  
GTGATGTTCTTCTGAGCCTTGCCA;

H2A/H2B F-TATTTGCTGTCTGGGTTAGGCGA, R-  
TTGCCTTTCCCTTCACTTTGCCAC;

H3 F-AGTCCGTCTTAAAGTCCTGAGCGA, R-  
AAAGAGCACCGAGCTTCTAATCCG

H4 F-TAGTGACTTTCGTGCTGTGCGTGT, R-CTTTCCCAAGCCTTTGCCTCCTTT;

H3/H4 F-TCTCTTTCACCGTCCACGATTGCT, R-  
TACGAGCCATCTCCGATTTGGGTT;

H3/H1 Intergenic F-ACACTGTCCCTTCAAACGCCTGTA, R-  
TCCTCGCCACATATGCATTACCGT;

Domino Exon6 F-GGCCGACGAGATGGGTCTG, R-GTAGGACGTGATACACACAT;

Domino poly(A)1 F-GCCAAAGCTGCCGAAGAAAGAAG, R-  
CACCCTGACTGCTGCTGATGAG.

Human

GAPDH 5' F-ACGTAGCTCAGGCCTCAAGA, R-AAGAAGATGCGGCTGACTGT

GAPDH 3' F-AACAAGGCCTTTTCCTCTCC, R-CCACTAACCAGTCAGCGTCA

H3F3A 5' F-CTTCCCTCCATTGTGTGTGATT, R-GGCAGACGGGGAAACTTAC

H3F3A 3' F-TGTTGTCTGTCTTCTGTGCTGTT, R-  
TTGCCACCAAGTCTATACTTTCC

HIST1H1C 5' F-AATGAAGAGCATGAAGCCCGAGG, R-  
AACTCGGGTACAAGTGGCAAGC

HIST1H1C 3' F-GCCAAAAGTGCTGCTAAGGCTGT, R-  
AAGTAACAGGGCAGAACAAGAAAAG

**Immunofluorescence.** Immunofluorescence was performed as described (Wagner et al., 2007) using anti-Mpm2 (Sigma) and anti-CstF50 primary antibodies.

## Results

The histone genes in *Drosophila* are clustered in a tandemly repeated unit containing one copy of each of the five genes (Fig. 8A). The 5' ends (start codons) of two pairs of divergently transcribed genes (H2A, H2B and H3, H4) are separated by 227 and 297 nt, respectively. The 3' ends of the H2A and H4 mRNAs are separated by 364 nt while, 257 nt separate the 3' ends of H2B and H1 mRNAs. There is a 1326 nt gap between the 3' end of the histone H3 mRNA and the start of histone H1. Thus, efficient processing and transcription termination are required to prevent transcription into neighboring ORFs.

Each of the histone genes contains canonical poly(A) sites downstream of the normal cleavage site which are present in multiple *Drosophila* species. Mutation of histone processing factors, SLBP (Lanzotti et al., 2002; Sullivan et al., 2001) or components of the U7 snRNP [(Godfrey et al., 2006) and submitted] results in the expression of polyadenylated histone mRNAs from each of the five histone genes. The presence of these polyadenylated mRNAs indicates that histone 3' end



processing is inefficient, allowing us to address the role of protein factors in this process.

**A subset of poly(A) factors are required for histone pre-mRNA processing in vivo.**

Three members of the poly(A) apparatus (CPSF73, CPSF100 and Symplekin) were previously shown to activate a histone misprocessing reporter (Wagner et al., 2007) but the effect of their knockdown on expression of the endogenous histone mRNAs had not been examined. To study *D. Melanogaster* histone pre-mRNA 3' end processing in vivo, Mindy Steiniger RNAi depleted both histone and general poly(A) processing machinery components in *D. Melanogaster* Dmel-2 cells, and performed an S1 nuclease protection assay to assess the processing of histone H2A mRNA (Fig. 8). Western and Northern blotting were used to assay the levels of processing factors (Fig. 9A) and their mRNAs (Fig. 10). Previous studies demonstrated the formation of several polyadenylated transcripts from the histone H2A gene in *Drosophila* mutants of U7 snRNA or SLBP (Godfrey et al., 2006; Lanzotti et al., 2002). Dmel-2 cells were bathed in dsRNAs targeting the transcripts of 3' end processing factors for five days and RNA and nuclear extracts (NEs) prepared from the cells. Total RNA was hybridized to a 670 nt probe. The 3' 340 nts of this probe base pair to the 5' end (including the processing site) of H2A. The probe also contained 310 nts complementary to the 3' flanking sequence which hybridized to polyadenylated histone H2A mRNAs (Lanzotti et al., 2002). 20 nts of non-histone sequence were added to the 5' end of the probe allowing us to distinguish any undigested probe from any transcripts that went past the end of the

probe (referred to as read-through transcripts). Hybridization of the probe was followed by digestion with S1 nuclease to map the 3' end of the RNA transcripts. Separation of the protected fragments by gel electrophoresis reveals properly processed histone mRNA as well as longer discrete fragments in samples depleted of various 3' end processing factors (Fig. 8C). Properly processed H2A mRNA yields a protected fragment of 340 nts, while misprocessed transcripts utilizing one of the downstream poly(A) sites resulted in several larger protected species between 340 and 650 nts (Fig. 8C, compare lanes 10 and 12).

Depletion of the histone specific pre-mRNA processing factor SLBP resulted in substantial misprocessing in vivo and multiple polyadenylated mRNA species (Fig. 8C, lane 10) as previously described for SLBP mutant embryos (Lanzotti et al., 2002) and larvae (Godfrey et al., 2006). Even when SLBP (and other factors) is depleted, the majority of the histone mRNA is still properly processed, indicating that the residual levels of the factors were sufficient for processing much of the histone mRNA, since in null mutants of SLBP or U7 snRNP only polyadenylated mRNA is produced (Godfrey et al., 2006).

In agreement with the results obtained with the reporter, knockdown of some components of the CPSF complex (CPSF73 or CPSF100) cause histone pre-mRNA misprocessing while knockdown of CPSF30 or CPSF160 had no effect on these transcripts (Fig. 8C, lanes 4-7). Symplekin depletion also results in H2A misprocessing (Fig. 8C, lane 8), while Fip knockdown had no detectable effect on endogenous H2A mRNA (Fig. 8C lane 9) though it scored weakly on the misprocessing reporter (Wagner et al., 2007). In contrast, depletion of CstF complex

components had no effect on histone 3' end formation (Fig. 8C lanes 1-3).

Interestingly, knockdown of polyadenylation factors primarily resulted in production of the long read-through product extending to the end of the probe (Fig. 8C, lanes 5, 6, 8).

To confirm that misprocessed H2A mRNA was polyadenylated, total RNA was fractionated on oligo d(T) cellulose to select polyA<sup>+</sup> RNA prior to the S1 nuclease protection assay. Enrichment of poly(A)<sup>+</sup> mRNA reveals that many of the misprocessed species are polyadenylated (Fig. 8D, lanes 5-8, 10-11). Depletion of both histone specific and general poly(A) factors resulted in a similar pattern of polyadenylated histone mRNAs in each sample. The pattern of protected fragments from histone H2A mRNA was strikingly different between total RNA and poly(A)<sup>+</sup> RNA of the CPSF73, CPSF100 and Symplekin knockdowns (compare Figs. 8C and 8D, lanes 5, 6, 8), while the pattern of protected fragments was similar in total and polyA<sup>+</sup> RNA in the SLBP and Lsm11 knockdowns (Figs. 8C and 9D, lanes 10 and 11). Most of the longest transcripts detected in depletion of the polyadenylation factors were not selected on dT cellulose and hence were not polyadenylated. Note that this protected fragment corresponds to the entire 310 histone 3' flanking sequence in the probe, and represents protection by all transcripts, regardless of their 3' ends, that extend past the end of the probe. The finding that in the CPSF73, CPSF100 and Symplekin knockdowns the majority of the misprocessed RNA is not polyadenylated, but rather represents an unprocessed read-through transcript, is consistent with a defect in polyadenylation as well as histone pre-mRNA processing when these factors are depleted. In contrast, depletion of histone specific factors

SLBP and Lsm11 resulted in similar ratios of polyadenylated H2A mRNA in both total and poly(A)<sup>+</sup> RNA, suggesting that there is little if any read-through unprocessed RNA accumulated. These data indicate that a subset of polyadenylation factors (CPSF73, CPSF100 and Symplekin) is required for proper histone pre-mRNA 3' end formation in vivo. Depletion of these factors results in accumulation of not only misprocessed polyadenylated histone mRNAs, as is seen in knockdowns of histone specific pre-mRNA processing factors, but also in unprocessed read-through transcripts, due to reduction in efficiency of polyadenylation.

### **Components of the core cleavage factor are codepleted by RNAi in vivo.**

Mindy next determined the levels of poly(A) and histone processing factors in each of the RNAi treated NEs by western blotting to confirm that knockdown was efficient. Strikingly, we found that depletion of some factors resulted in codepletion of other factors, suggesting that they are present in a complex. Similar results have been observed for some homologous polyadenylation factors in yeast (Kyburz et al., 2003; Zhao et al., 1999b). Depletion of CPSF73, CPSF100 or CPSF160 results in codepletion of CPSF73, CPSF100 and Symplekin, while CSPF30 depletion has only a mild affect on the level of CPSF100 (Fig. 9A, lanes 1-3 and 10). Additionally, knockdown of Symplekin also causes codepletion of CPSF complex components CPSF73 and CPSF100 (Fig. 9A lane 4). Note that there are not antibodies available for all the factors that we knocked down. Depletion of the CstF subunits CstF64 and CstF77, resulted in codepletion of CstF50 (Fig. 9A, lanes 12 and 13), although knockdown of CstF50 showed only modest codepletion of CstF77 (Fig. 9A, lane 7).

There was no effect on the histone specific factors SLBP and Lsm11 with knockdown of the polyadenylation factors with the lone exception of a variable SLBP depletion in CPSF30 knockdowns, which was not sufficient to affect histone mRNA processing. Knockdown of SLBP had a minor affect on CPSF100 level while Lsm11 knockdown did not affect any other factors tested (Fig. 9A, lanes 5-6).

Notably, while knockdown of CPSF160 also resulted in reduction in the levels of CPSF73, CPSF100 and Symplekin, CPSF160 depletion does not have an effect on histone pre-mRNA processing ((Wagner et al., 2007) and Figs. 8C and 8D, lane 7). Similarly, knockdown of CstF64 resulted in codepletion of Symplekin (Fig. 2A, lane 12), but also did not affect histone pre-mRNA processing (Figs. 8C and 8D, lane 2). None of the knockdowns significantly affect mRNA levels of other factors, indicating that the codepletions are not due to the reduction of levels of specific mRNAs (Figure 10). These data indicate that components of both the CPSF and CstF complexes can affect the stability of proteins within their respective complexes. Additionally, Symplekin protein levels are reduced by knockdown of the CPSF components CPSF73, CPSF100 and CPSF160 as well as CstF64 but not the other CstF subunits.

### **CPSF73, CPSF100 and Symplekin form a stable complex and also interact with histone mRNA specific processing factors**

To directly test for protein-protein interactions among *Drosophila* 3' end processing factors, colP experiments were performed with NEs from Dmel-2 cells. Mindy prepared polyclonal CPSF73 and Symplekin antibodies which were effective

for IP, and used these to characterize the interactions between endogenous Symplekin and the members of the CPSF complex. CPSF100 and Symplekin colPs with CPSF73 under stringent conditions (50 mM Tris, pH 8.0, 1% NP40, 0.25% deoxycholate, 150 mM NaCl, Fig. 9B, lane 4). This result was confirmed by colP of CPSF73 and CPSF100 with the  $\alpha$ -Symplekin antibody (Fig. 9B, lane 8). The high efficiency of colP of these three endogenous factors strongly supports the idea that CPSF73, CPSF100 and Symplekin are part of a stable complex(es) in vivo.

To determine if components of the cleavage factor interact with specific histone pre-mRNA processing factors, colP experiments with  $\alpha$ -SLBP and  $\alpha$ -Lsm11 antibodies were performed using NE from Dmel-2 cells. CPSF73 colPs with both SLBP and Lsm11 (Fig. 9C, lanes 3-5) under less stringent conditions (50mM Tris, pH 8.0, 0.1% NP-40, 150 mM NaCl), although neither interaction is as strong as that between CPSF73 and Symplekin as these interactions are undetectable under stringent conditions (data not shown). Symplekin also colPs with SLBP and, to a lesser extent, Lsm11 (Fig. 9C, lanes 3-5). While over 50% of the Symplekin, CPSF73 and CPSF100 coimmunoprecipitated with one another, only about 1% of the CPSF73 or Symplekin precipitated with SLBP or Lsm11. CstF77 did not colP with Symplekin, CPSF73, Lsm11 or SLBP, consistent with depletion of CstF77 not having an effect on the concentrations of these proteins. These interactions are not RNA dependent as neither RNase treatment of the IPs nor actinomycin D treatment of the cells prior to nuclear extraction prevented coimmunoprecipitation (Fig. 11).

Collectively, these data point to a core complex of cleavage factor components consisting of CPSF73, CPSF100 and Symplekin which interact very

strongly with one another. This core complex may then bind different accessory proteins depending on the type of pre-mRNA to be processed. A schematic of possible interactions is shown in Figure 9D, indicating how a common core of CPSF73, CPSF100 and Symplekin might interact to form two different cleavage factors for processing of poly(A) mRNAs (Fig. 9D, left) versus histone mRNAs (Fig. 9D, right).

### **Components of the cleavage factor associate with histone genes in vivo**

RNA processing factors have been shown via ChIP to associate with genes (Komarnitsky et al., 2000; Licatalosi et al., 2002a; Proudfoot, 2004; Swinburne et al., 2006), either as complexes bound to the nascent RNA or by direct association with RNA Pol II. I investigated whether any of these factors associate with the *Drosophila* histone genes in vivo. Primer sets located within the coding regions of four of the genes (H2A, H2B, H3, and H4) as well as three noncoding regions, including the promoter regions of H2A/H2B and H3/H4 plus a region downstream of the H3 gene in the largest intergenic region contained in the repeat were used to amplify regions of the ~5kb histone gene cluster (Fig. 12A). There are multiple polyadenylation sites in this 1.4 kb intergenic region that are only used when histone mRNA processing is affected (Lanzotti et al., 2002), and the probe is located 472 to 657 nts after the histone processing site and 669 nts 5' of the histone H1 start codon. It should be noted that the small size of the histone genes (~400 bp) approaches the limit of the resolution of the chromatin shearing required for the ChIP assay (200-500 bp fragments) and thus differences between the 5' and 3' ends of these genes may not be readily distinguishable.

As a control for the association of canonical poly(A) factors with a gene encoding a polyadenylated mRNA, primers were also designed to exon 6 and the first of two poly(A) sites located in the domino gene (Fig. 12A). In contrast to the histone genes, these probes are well separated. ChIP experiments were performed with antibodies to CPSF73, CstF50, RNA Pol II (Rpb3) and Symplekin. The relative amounts of DNA associated with each fragment were quantified using qPCR and SYBRGreen fluorescence. Using an antibody which recognizes the Rpb3 subunit of RNA Pol II and hence all forms of RNA Pol II without preference for phosphorylation state of the CTD, I found that RNA Pol II associates with all portions of the histone genes and the domino gene. Within the promoter and coding regions of each of the genes, the polymerase occupancy is between 40% and 80% of maximum whereas there is no RNA Pol II present in the intergenic region between the H3 and H1 genes (Fig. 3B, blue bars).

I next examined the localization of two components of the cleavage factor, CPSF73 and Symplekin. The domino gene shows an increased pattern of recruitment of both factors at the 3' end of the gene relative to exon 6 (Figure 12B, red and green bars). Both CPSF73 and Symplekin are localized throughout the histone gene cluster with peak levels similar to those found at the 3' end of the domino gene. Again, essentially no Symplekin or CPSF73 was detected in the intergenic between histone H3 and histone H1 in control cells. As a control, we carried out ChIP analysis with CstF50, a factor which has no role in histone pre-mRNA processing. CstF50 showed strong localization to the 3' end of the domino gene and a much lower level of association with the histone genes than CPSF73 or



Symplekin. Not surprisingly, there is little localization of CstF50 with the H3/H1 intergenic region (Fig. 12B, purple bars). Thus CPSF73 and Symplekin, factors involved in histone mRNA maturation, display cotranscriptional recruitment to histone genes as well as a canonical poly(A) gene.

### **Mammalian SLBP associates with histone genes in vivo.**

To determine whether histone specific processing factors associate with histone genes in vivo, I wanted to perform ChIP experiments with an antibody recognizing either SLBP or a U7 snRNP component. As our dSLBP and dLsm11 antibodies gave high background signals in my ChIP assay, I elected to carry out ChIP experiments in HeLa cells stably expressing HA-tagged SLBP. This allowed me to IP with both an HA antibody as well as an antibody which recognizes endogenous SLBP. I designed oligonucleotide sets to amplify the 5' and 3' ends of three genes; GAPDH, H3F3A, which encodes a polyadenylated H3.3 mRNA, and HIST1H1C, which encodes a canonical replication dependent histone mRNA. Both the HA and SLBP antibodies gave very low signals on the GAPDH and H3F3A genes which I concluded to be background levels, however, both gave significant signals on the HIST1H1C gene. The HA antibody gave lower signals at both ends of the gene (not much above background), but showed a 2-fold increase from the 5' to the 3' end of the gene, consistent with cotranscriptional recruitment of SLBP (Fig. 13, black bars). These low levels of ChIP signal may be due to the fact that the HA-SLBP is overexpressed and may not be incorporated into active complexes at the same frequency as endogenous protein. ChIP experiments with the SLBP antibody gave signals well above background on both the 5' and 3' ends of the HIST1H1C

gene and showed a similar 2-fold increase in signal at the 3' end of the gene (Fig. 13, purple bars). These data are consistent with SLBP being recruited to histone genes cotranscriptionally.

### **SLBP knockdown causes recruitment of additional poly(A) factors to histone genes**

Since knockdown of SLBP resulted in misprocessing and polyadenylation of histone mRNAs (Fig. 8C, lane10) I investigated how recruitment of the processing machinery to the histone genes was affected by depletion of SLBP. ChIP lysates were prepared from Dmel-2 cells in which SLBP was depleted by RNAi and IPs were performed as in previous experiments. In these knockdown experiments I focused on the H3 gene using the primer sets which correspond to the 5' end and promoter region (primer set H3/H4), the 3' end of the coding region (primer set H3) and the region between the H3 and H1 genes which contains the two distal poly(A) sites (primer set H3/H1 intergenic). The domino poly(A) site was used as a control. RNA Pol II drops from ~60% maximal occupancy at the 3' end of histone H3 genes in control cells to ~40% in this same region in SLBP knockdown cells (Fig. 14A, top left). At the same time, RNA Pol II occupancy in the intergenic region increases from less than 10% in control cells to >40% in SLBP knockdown cells. CPSF73 and Symplekin show similar recruitment patterns at the promoter and 3' end in control and SLBP knockdown samples, however, both of these proteins also display a dramatic increase in ChIP signal in the intergenic region in SLBP knockdown cells (compare Fig. 14A, top right and bottom left). The most remarkable difference between control and SLBP knockdown cells is the enrichment of CstF50 across the

histone H3 gene (Fig. 14A, bottom right). In control cells, this protein is recruited at only a low level, however, in cells depleted of SLBP it shows occupancy of >50% of maximum at the 3' end and intergenic regions. The intergenic region which contains two distal poly(A) sites experiences an increase in CstF50 from ~5% to ~70% in SLBP knockdown cells. This is consistent with recruitment of CstF50 to histone genes as part of the polyadenylation of histone pre-mRNAs in the absence of SLBP.

Immunofluorescence experiments also revealed an increased recruitment of CstF50 to HLBs when SLBP is RNAi-depleted (Figure 14B). Cells were analyzed by immunofluorescence for Mpm2, an antigen that localizes to HLBs (White et al., 2007) and hence histone genes in S-phase *Drosophila* cells (>50% of cells), and CstF50. CstF50 is generally localized throughout the nucleoplasm in control cells, consistent with a role in processing a large number of genes. Foci of CstF50 were clearly visible in some of the SLBP knockdown cells and in 64% of the cells these foci colocalized with Mpm2, consistent with them being in the HLB. In control cells, CstF50 foci were generally not visible, although in 34% of cells there were significant amounts of CstF50 that overlapped in localization with Mpm2 (Fig. 14B). Thus SLBP knockdown results in an increased association of CstF50 with the HLB.

To determine the effect of simultaneously depleting both a histone and a poly(A) factor on histone pre-mRNA processing, protein lysates and RNA were made from Dmel-2 cells depleted of both SLBP and CstF50. Efficient knockdown of both proteins in the double knockdown was confirmed by western blot (Fig. 14C). An S1 nuclease protection assay was then performed on the RNA from the double knockdown as well as on RNA from SLBP and CstF50 individual knockdowns. SLBP

knockdown again resulted in substantial misprocessing and polyadenylation (Fig. 14D, lanes 2 and 6) while depletion of CstF50 had no effect on histone pre-mRNA processing (Fig. 14D, lanes 3 and 7). The double knockdown, however, shows an intermediate phenotype in which the distal poly(A) site is preferentially used for polyadenylation (Fig. 14D, lanes 4 and 8), consistent with both histone pre-mRNA processing being affected by SLBP depletion and polyadenylation being affected by CstF50 depletion. Collectively, the data in Figure 4 demonstrate a role for the CstF complex in polyadenylation of misprocessed histone mRNAs and that CstF50 is not required for normal processing.

#### **Depletion of Symplekin by RNAi-depletion of CPSF160 or CstF64 does not affect recruitment of the cleavage factor to histone genes**

Knockdown of CstF64 or CPSF160 results in codepletion of factors required for histone pre-mRNA processing yet does not result in polyadenylation of histone mRNAs (Fig. 8C, lanes 2 and 7). Knockdown of CPSF160 codepletes CPSF73, CPSF100 and Symplekin, while CstF64 knockdown results in depletion of Symplekin (Fig. 9A, lanes 10 and 12). To determine whether this codepletion affects association of common cleavage factor components with either histone or poly(A) genes, I prepared lysates from CstF64 and CPSF160 knockdown cells and performed ChIP. ChIP profiles on histone genes in the CstF64 knockdown cells showed a similar pattern as in control cells. RNA Pol II ChIP profiles showed maximum occupancy at the 3' end of the H3 gene with no significant association in the intergenic region between H3 and H1. Also, RNA Pol II had approximately 60% of maximal association with the promoter region of H3 and the proximal domino

poly(A) site (Fig. 15A, top left). These numbers reveal a slightly increased association of RNA Pol II on both ends of the histone gene relative to the domino gene. CPSF73 is associated with both the 5' and 3' ends of histone genes (Fig. 15A, top right). Symplekin shows a similar pattern of association with the histone gene as both RNA Pol II and CPSF73, although there is a substantial decrease in the relative amount of Symplekin associated with the domino poly(A) site (Fig. 15A, bottom left). The reduction in Symplekin at the 3' end of the domino gene with no effect on its association with the histone genes, is striking. This is consistent with CstF64 knockdown resulting in codepletion of Symplekin but with no effect on histone mRNA processing.

I next examined the effect of CPSF160 knockdown on association of RNA Pol II and RNA processing factors with the histone H3 gene and the proximal domino poly(A) site. Similar to the CstF64 knockdown, there is a slight increase of RNA Pol II association with the histone gene while there is a modest decrease in its association with the domino 3' end (Fig. 15B, top left). CPSF73, one of the factors codepleted by CPSF160 knockdown shows a modest decrease in association with the 5' end of histone H3 gene but a high level of association with the 3' end of histone H3 gene (Fig. 15B). Strikingly, CPSF73 levels are down to 10% of maximum value in CPSF160 knockdown cells on the poly(A) site of the domino gene (Fig. 15B, top right). Symplekin profiling reveals that its association with the histone genes is slightly higher in the CSPF160 knockdown than in wild type cells, but that binding to the proximal domino poly(A) site decreases (Fig. 15B, bottom left). These data show that codepletion of Symplekin and CPSF73 by knockdown of

polyadenylation factors which are not required for histone 3' end processing, does not affect recruitment of CPSF73 or Symplekin to histone genes, but it does affect recruitment to the domino gene. Together these data suggest that there are two distinct cleavage complexes and that knockdown of a factor which is not required for histone pre-mRNA processing (not present in the histone cleavage complex) may codeplete the polyadenylation cleavage complex, but not the histone cleavage complex.

### **SLBP is in a complex with RNA Pol II**

As many RNA processing factors are recruited to nascent RNA cotranscriptionally by associating with the CTD of RNA Pol II (Proudfoot et al., 2002), I wanted to test whether SLBP associates with the polymerase. I performed IP experiments using antibodies against Ser2P CTD, Ser5P CTD or Rpb3 and looked for coIP of SLBP, Symplekin or CPSF73. To determine whether there was any interaction between RNA Pol II and SLBP, I decided to perform IP experiments on formaldehyde crosslinked cell lysates under conditions similar to those used for the ChIP experiments and these experiments revealed an interesting result. I began by looking for association of Symplekin and CPSF73 with the polymerase as these are components of the general polyadenylation machinery and I believed it more likely that I would find interactions between these proteins and RNA Pol II. As expected, both Symplekin and CPSF73 interact with the Ser2P CTD which is the modification associated with the 3' ends of genes (Fig. 16A, lane 4). Both proteins also precipitate with the Rpb3 antibody which has no preference for the phosphorylation state of the CTD (Fig. 16A, lane 3), but only Symplekin shows an

association with the Ser5P form of RNA Pol II indicating that this protein may associate with the 5' ends of genes (Fig. 16A, lane 5). Lanes 6 and 7 of figure 16A demonstrate that both CPSF73 and Symplekin efficiently IP under crosslinked conditions. It should be noted that the coIP of Symplekin and CPSF73 with the myc antibody is aleatoric as SLBP does not coIP with myc and in the vast majority of experiments performed, neither does Symplekin or CPSF73.

In crosslinked lysates, SLBP coIPs with the Rpb3 antibody, but not the Ser2P or Ser5P phosphospecific antibodies (Figure 16A, lanes 3-5); indicating that this association, whether direct or indirect, is not dependent on either of the well characterized modifications of the CTD. Interestingly, we are also able to pull down SLBP with Symplekin under crosslinked conditions, an interaction we are unable to observe in NEs, though we can readily detect the reciprocal interaction (Figure 16, lane 7). These data indicate that SLBP and RNA Pol II exist in a complex together.

To determine whether or not associations between RNA Pol II and 3' end processing factors was dependent upon active transcription, cells were treated with actinomycin D, a general transcription inhibitor, and NEs prepared. IP of Symplekin and CstF77 demonstrated that association of these factors with RNA Pol II was, indeed, dependent on active transcription as actinomycin D treatment abrogated these interactions (Fig. 16B, lanes 3-4). The interaction between SLBP and RNA Pol II, however, was not dependent on active transcription (Fig. 16B, lanes 3-4), indicating that this association may be due to high concentrations of RNA Pol II and SLBP in HLBs. The significance of this association is a subject worthy of further study.

## Discussion

All mRNAs in eukaryotic cells are transcribed by RNA polymerase II, and the first step in 3' end formation of all mRNAs is endonucleolytic cleavage. Cleavage is coupled with polyadenylation for most genes, but cleavage is the only step in 3' end formation of the metazoan replication-dependent histone mRNAs. In the last few years, it has become apparent that there are factors shared between these two processes, and that CPSF73 catalyzes the endonucleolytic cleavage of both polyadenylated mRNAs (Mandel et al., 2006; Ryan et al., 2004) and histone mRNAs (Dominski et al., 2005b). Here Mindy and I provide evidence that a subset of poly(A) factors, CPSF73, CPSF100 and Symplekin, which are involved in both polyadenylation and histone pre-mRNA processing, are present in a complex in *Drosophila* cells.

### **CPSF73, CPSF100 and Symplekin comprise the mRNA processing core cleavage factor.**

Depletion of the cleavage and poly(A) factors CPSF73, CPSF100 and Symplekin from *Drosophila* cultured cells results in misprocessing of histone pre-mRNA. In contrast to depletion of SLBP or U7 snRNP where most of the read-through mRNAs are polyadenylated, when CPSF73, CPSF100 or Symplekin are depleted the majority of misprocessed histone mRNAs are run-on transcripts that are not polyadenylated (Fig. 8). These data are consistent with knockdown of each of these three factors also causing defects in polyadenylation in vivo (Wagner et al., 2007). These experiments confirm that only a subset of poly(A) components are required for proper histone pre-mRNA processing in vivo.



A striking finding is that depletion of some proteins using RNAi results in codepletion of other components (Fig. 9). Since RNAi does not significantly affect the mRNA levels of other factors (Fig. 10), a likely interpretation of this result is that the codepleted factors are present in a complex and that reduction of a single component results in degradation of the other components. For example, depletion of CPSF73 causes codepletion of CPSF100 and Symplekin, and similarly, depletion of CPSF100 or Symplekin results in depletion all three components. Depletion of these factors does not result in codepletion of SLBP, CstF50 or CstF77. Depletion of a U7 specific protein, Lsm11, does not affect the levels of any other factors we tested, although it does reduce the level of Lsm10 (Wagner et al., 2007). These results are consistent with the interpretation that there is a stable complex(es) containing CPSF73, CPSF100 and Symplekin. Additionally, a second complex exists containing, Cst50, Cst64 and Cst77, in agreement with the biochemical studies in mammalian cells of Manley and coworkers (Hatton et al., 2000; Takagaki et al., 1990).

The presence of a stable complex containing CPSF73, CPSF100 and Symplekin is further supported by IP experiments using either  $\alpha$ -CPSF73 or  $\alpha$ -Symplekin antibodies in stringent conditions (buffer containing an ionic detergent). A substantial fraction (more than 50%) of the endogenous CPSF73, CPSF100 and Symplekin in NEs coprecipitate with one another (Fig. 9). Thus these three proteins are part of a stable complex(es) in vivo. I observe much less efficient CoIP between Symplekin or CPSF73 with SLBP or Lsm11, although there is also significant coimmunoprecipitation of these endogenous factors. I propose that the core

cleavage factor is recruited to the histone pre-mRNA via interactions with histone-specific proteins SLBP and Lsm11. At any one time only a small fraction of the cleavage factor is associated with the other histone processing factors. It is likely that there are other components of the histone processing factor which have not yet been identified, given the large number of polypeptides required for cleavage of polyadenylated mRNAs (Mandel et al., 2007).

**The core cleavage factor is associated with histone genes in vivo.**

To determine whether processing factors are associated with nascent histone pre-mRNAs, ChIP assays were performed using RNA Pol II, histone pre-mRNA processing factors and polyadenylation factors. These experiments revealed that under normal conditions RNA Pol II does not proceed more than 472 nts 3' end of the histone mRNA, suggesting that processing occurs shortly after transcription of the HDE. These results agree with the in vitro results of Adamson and Price (Adamson and Price, 2003) who detected strong pause sites just downstream of the HDE of all 5 *Drosophila* histone genes. Pausing of RNA Pol II at these sites could allow time for the processing complex to assemble and subsequent transcription termination. I also detected association of histone pre-mRNA processing factors CPSF73 and Symplekin with the 3' end of histone genes in vivo, consistent with formation of the processing complex on genes. ChIP experiments with SLBP and Lsm11 antibodies were inconclusive due to high background levels. These data suggest that the cleavage complex is recruited cotranscriptionally to the histone pre-mRNA.

I did not observe large amounts of CstF50 associated with the histone genes under normal conditions. However when SLBP was knocked down, RNA Pol II was detected 3' of the normal processing site, and in addition to CPSF73 and Symplekin, CstF50 was also associated with the histone genes, consistent with its role in polyadenylation. Finally, ChIP studies on lysates in which CPSF160 or CstF64 are knocked down show selective depletion of Symplekin and/or CPSF73 from poly(A) genes, with no affect on their recruitment to histone genes, supporting the notion of distinct cleavage complexes for histone and poly(A) genes.

**Two distinct complexes, both containing the core cleavage factor, are required for histone mRNA and poly(A) mRNA 3' end formation.**

Even though knockdown of either CPSF160 or CstF64, proteins only required for polyadenylation, reduces the level of Symplekin, the association of Symplekin with the 3' end of histone genes was not affected under these conditions. Similarly, knockdown of CPSF160 codepletes CPSF73 and CPSF100, yet association of CPSF73 with the 3' end of histone genes was unaffected in a CPSF160 knockdown (Figure 15). However, there was a reduction in the association of Symplekin and CPSF73 with the 3' end of the domino gene. These data are consistent with the CPSF complex, which contains CPSF160, being one of the stable complexes containing Symplekin. Symplekin and CPSF73 continue to associate with the histone genes in the CPSF160 and CstF64 knockdowns, and in vivo histone RNA processing is not affected, suggesting that a distinct cleavage factor is recruited to the histone pre-mRNAs. ChIP experiments also show that there was no RNA Pol II associated in the intergenic region, consistent with the failure to detect read-through

transcripts in vivo. Collectively, these results support our hypothesis that there are distinct cleavage complexes for these two classes of genes, and that one complex may be depleted without affecting the levels or functionality of the other.

Interestingly, a comparison of histone and poly(A) pre-mRNA 3' cis elements provides support to this model. Histone pre-mRNAs have two important cis elements, the highly conserved stem loop (SL) which binds the SLBP and a less conserved histone downstream element (HDE) which base pairs to U7 snRNA in the U7 snRNP. The SL and the HDE are analogous to the highly conserved AAUAAA hexamer and the less conserved GU-rich downstream element (DSE) in the 3' UTRs of canonical poly(A) mRNAs. The AAUAAA hexamer interacts with CPSF160 while CstF64 binds the DSE. In both cases there is endonucleolytic cleavage between these elements due to recruitment and/or activation of a cleavage factor containing CPSF73 and CPSF100 as well as Symplekin. Thus there is a common core of the cleavage factor that is utilized for maturation of both histone and poly (A) mRNAs.

It is likely that the cell physically maintains these two separate complexes, and directs them to the proper targets. For polyadenylation, CPSF160, a component of the polyA specific machinery directly binds the AAUAAA and thus brings CPSF73 and CPSF100 to the pre-mRNA, albeit in an inactive form, since cleavage requires assembly of additional factors. For histone mRNA, recruitment of the cleavage factor components is the last step in processing. Concentration of the U7 snRNP in the HLB near the histone genes is critical for efficient processing of histone mRNA in vivo (Wagner et al., 2007). Possibly a U7 snRNP component or an as yet unknown

protein, with a high local concentration in the HLB can recruit the histone cleavage factor to the HLB.

### **Histone pre-mRNA 3' end processing is cotranscriptional.**

The factors involved in transcription of the histone genes include many proteins required for transcription of canonical polyadenylated mRNAs. Unlike mammalian snRNAs, there is no evidence that formation of the histone mRNA 3' end is affected by the promoter used for transcription (Pilch and Marzluff, 1991; Whitelaw et al., 1986). A histone reporter driven by the *Drosophila* actin promoter forms histone 3' ends just as effectively as the same reporter driven by the endogenous histone promoter (B. Burch and W.F.M., unpublished). Also a histone promoter driving a polyadenylated mRNA results in efficient polyadenylation (B. Burch, E.J. Wagner and W.F.M., unpublished). It is possible that sequences in the coding region of the histone mRNA also affects processing efficiency in vivo. The ~400 nt coding regions of histone genes are highly conserved (Debry and Marzluff, 1994) and there are factors that can bind to the nascent transcript that could help recruit processing factors (Friend et al., 2007). Thus, I believe that a major factor in efficient histone pre-mRNA processing is a high local concentration of the processing factors at the HLB together with the intrinsic pause site after the HDE which would allow time for assembly of the processing complex.

Many factors that participate in transcription and processing of canonical poly(A) mRNAs via interaction with the CTD [P-TEFb (Taube et al., 2002), capping enzymes (Schroeder et al., 2000)] are likely also recruited to the histone genes.

There is substantial evidence that some components of the polyadenylation machinery can bind to the CTD of RNA Pol II [reviewed in (Buratowski, 2005a; Neugebauer, 2002)], possibly shortly after initiation of transcription, and coordinate transcription with 3' end processing. CstF50, CstF64 and CstF77 interact directly with the CTD, with a preference for the phosphorylated CTD, possibly to facilitate 3' end processing of poly(A) transcripts in mammals (McCracken et al., 1997) and yeast (Barilla et al., 2001). The yeast CPSF160 orthologue can also bind directly to the CTD (Dichtl et al., 2002). The interactions of CstF with the CTD may account for our observation that a factor not involved in histone 3' end processing, CstF50, has 20-25% occupancy on the histone genes. Moreover, recent studies have suggested that there is substantial difference in the modifications on RNA Pol II in Cajal bodies (Xie and Pombo, 2006) which may account for the fact that we cannot detect interactions between SLBP and RNA Pol II with Ser2P or Ser5P CTD, but we do detect this interaction using an antibody which recognizes a subunit of RNA Pol II which is independent of the CTD. While I believe that SLBP is in a complex with RNA Pol II, I currently have no functional data for this interaction and it could be a result of high concentration of these proteins at histone genes. Therefore, I currently favor a model of a high local concentration of processing factors in HLBs rather than direct recruitment by the CTD of RNA Pol II.

#### **D. melanogaster histone mRNA 3' end processing.**

The histone genes are localized near the histone locus body (HLB), a site at which U7 snRNP but not SLBP is concentrated. There is also a locally high concentration of Symplekin in the HLB [(Wagner et al., 2007); D. Tatomer., K.D.S,

M.S. and W.F.M., unpublished]. As histone genes are transcribed, SLBP binds the SL and U7 snRNP associates with the HDE. RNA Pol II pauses after the HDE is transcribed (Adamson and Price, 2003), allowing time for the SLBP and U7 snRNP complex to assemble on the histone pre-mRNA. The cleavage factor, containing at least Symplekin, CPSF73, CPSF100 and possibly other factors, associates with the SLBP/U7snRNP/pre-mRNA complex and cleaves the pre-mRNA (Figure 17A).

Under normal conditions I do not observe RNA Pol II downstream of the cleavage site, suggesting that transcription termination occurs rapidly after processing, as has been suggested for mammalian histone genes (Chodchoy et al., 1991) . It is possible that CPSF73 rapidly degrades the 3' cleavage product after processing (Yang et al., 2009) promoting transcription termination. If there is a deficiency in any of these factors, processing does not occur while the RNA Pol II is paused and RNA Pol II may eventually resume transcription and transcribe the polyadenylation sites. If the deficiency is in a protein only required for efficient histone pre-mRNA processing (SLBP or Lsm11), additional factors for poly(A) pre-mRNA processing will be recruited and the histone pre-mRNA will be cleaved and polyadenylated (Figure 17B). If the deficiency is in a protein necessary for both histone processing and processing of poly(A) transcripts, RNA Pol II will continue to transcribe and only small amounts of polyadenylation may occur. The result is read-through transcripts which are likely rapidly degraded.

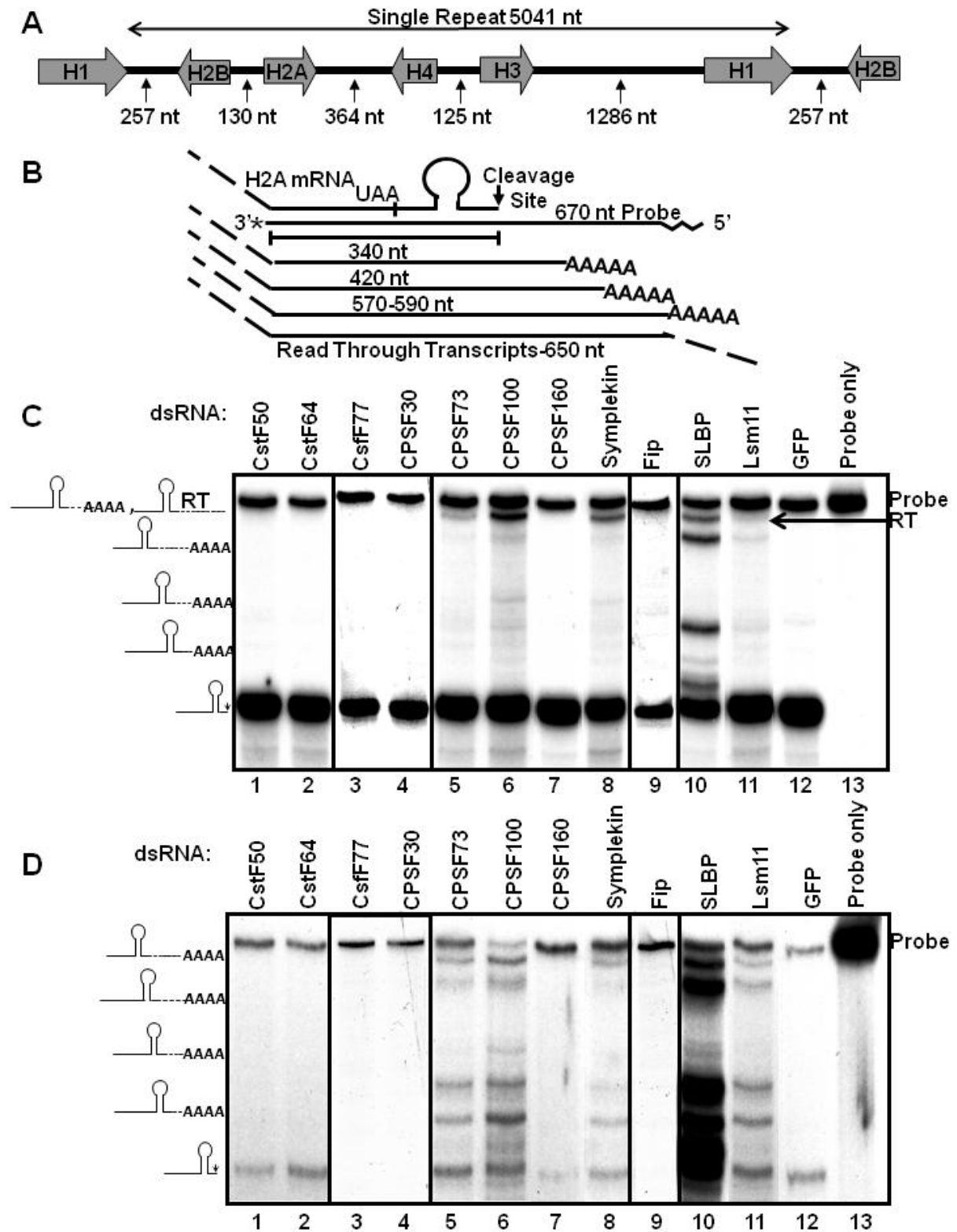
Read-through transcription also occurs when NELF (a negative elongation factor) is depleted in *Drosophila* (Gilchrist et al., 2008; Wagner et al., 2007) and human cells (Narita et al., 2007). This may be due to inefficient pausing of RNA Pol

II resulting in inefficient histone 3' end formation. Read-through transcription also occurs when the U7 snRNP is not concentrated at the HLB (Wagner et al., 2007), probably as a result of inefficient recruitment of U7 snRNP while RNA Pol II is paused. Thus, efficient processing in vivo requires the proper concentration of processing factors at the histone locus, pausing by RNA Pol II and likely other factors involved in coordinating processing and transcription. Deficiencies in any of these processes cause read-through past the normal termination site and subsequent polyadenylation of many of the read-through transcripts. The molecular details of how these multiple pathways intersect at the processing site to accomplish efficient processing and transcription termination remain to be elucidated.

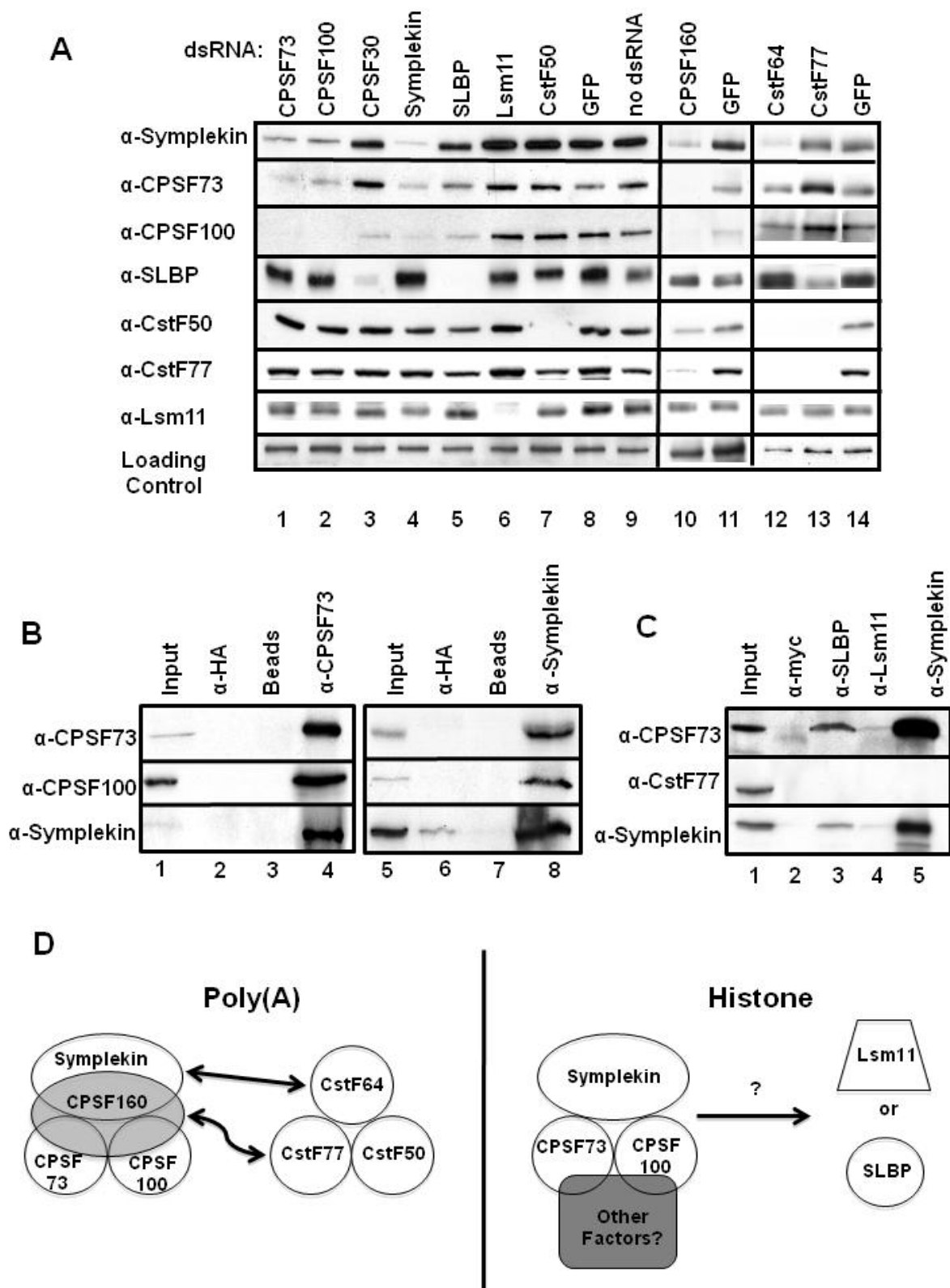


**Figure 8. Knockdown of pre-mRNA processing factors results in**

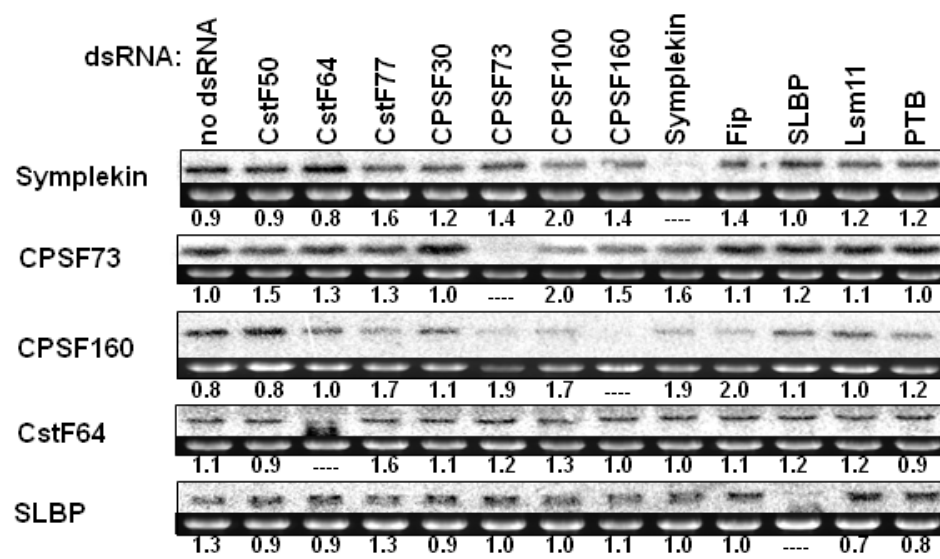
**misprocessed histone mRNA.** (A) A schematic of one of the tandem repeat units of the histone gene locus is shown. The numbers indicate the distance between the 5' or 3' ends of the mRNAs. The distance between the H3 and H1 genes is the distance between the 3' end of the H3 mRNA and the translation start codon of the histone H1 gene. (B) Schematic of S1 nuclease protection assay. A 670 nt DNA fragment containing 650 nts of the 3' end of H2A gene and 20 nts of plasmid sequence at the 5' end was labeled with  $\gamma$ - $^{32}\text{P}$ -dCTP on the 3' end. The probe was hybridized to total cell RNA and digested with S1 nuclease. Properly processed H2A mRNA yields a protected fragment of 340 nt extending to the 3' end of the mRNA while misprocessed transcripts give fragments ranging in size from 340-600nt. Any read-through protect a 650 nt fragment. (C) Dmel-2 cells were treated with dsRNAs targeting the indicated proteins above each lane for 5 days and total RNA was harvested by Trizol extraction. 12.5  $\mu\text{g}$  of total RNA was subjected to S1 nuclease protection assay and protected fragments were resolved on a 6% acrylamide gel. Diagrams at left indicate RNA species corresponding to protected fragments. The undigested probe is shown in lane 13. Read-through transcripts (RT) are marked with an arrow. (D) 25  $\mu\text{g}$  of total RNA was purified on oligo(dT) cellulose and the purified poly(A)<sup>+</sup> RNA was used for S1 nuclease protection assay as in (B). The undigested probe is shown in lane 13. A small amount of properly processed mRNA nonspecifically binds to the oligo(dT) cellulose.



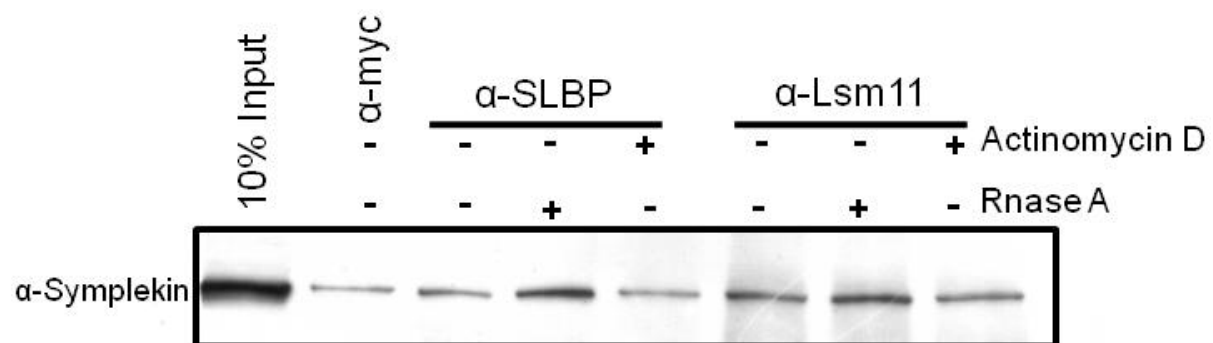
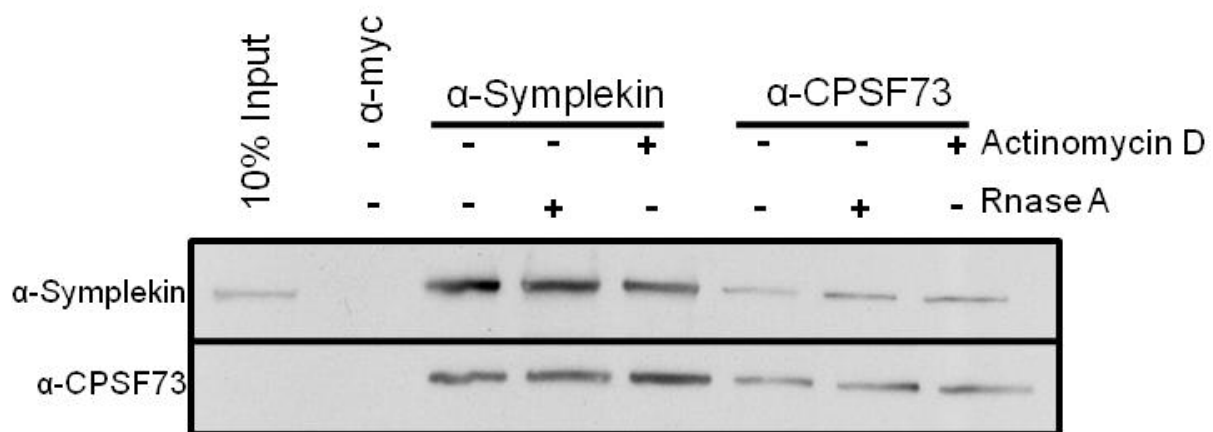
**Figure 9. The core cleavage factor interacts with histone specific processing factors.** (A) Dmel-2 cells were treated with the indicated dsRNA for 5 days and NEs were prepared. 15 µg of NE were resolved by SDS-PAGE and western blots were performed using indicated antibodies. The loading control is a cross-reacting band from the Symplekin blot. (B) 500 µg of Dmel-2 NE was used in IP experiments with indicated antibodies under RIPA (high detergent) conditions. Precipitates were run on SDS-PAGE gels and western blots performed as in A. α-HA and beads alone IPs were included as controls. Inputs represent 2.5% of starting material. (C) IPs were performed as in A except under low detergent conditions. Western blots performed as in A. (D) Diagram of core cleavage factor interactions with poly(A) and histone specific RNA processing factors.



**Figure 10. Codepletion is not due to reduced mRNA levels.** 10 ug of total RNA from knockdown cells was resolved on a 1.5% Formaldehyde agarose gel and transferred to Hybond-N<sup>+</sup> membrane. Probes for CPSF73, 100 and 160, CstF64 and SLBP were used to assess mRNA levels, top panels, and values were normalized to ribosomal RNA, bottom panels.



**Figure 11. Interactions between cleavage factor and histone pre-mRNA processing factors are not RNA-dependent.** IPs were performed as in figure 9 from untreated NEs with or without 30 ug/mL RNAase A, or from NEs treated with actinomycin D (1ug/mL final concentration) for 30 min.

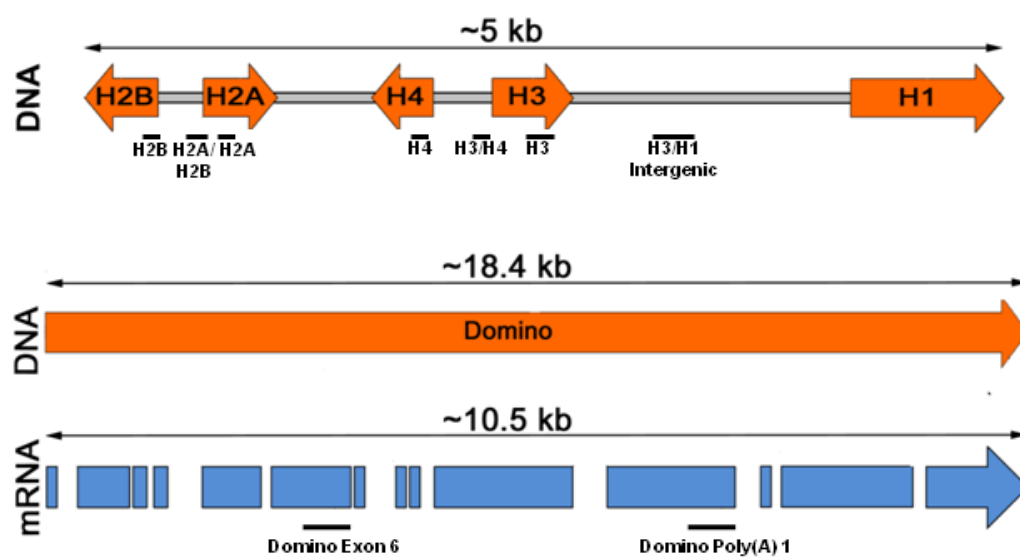




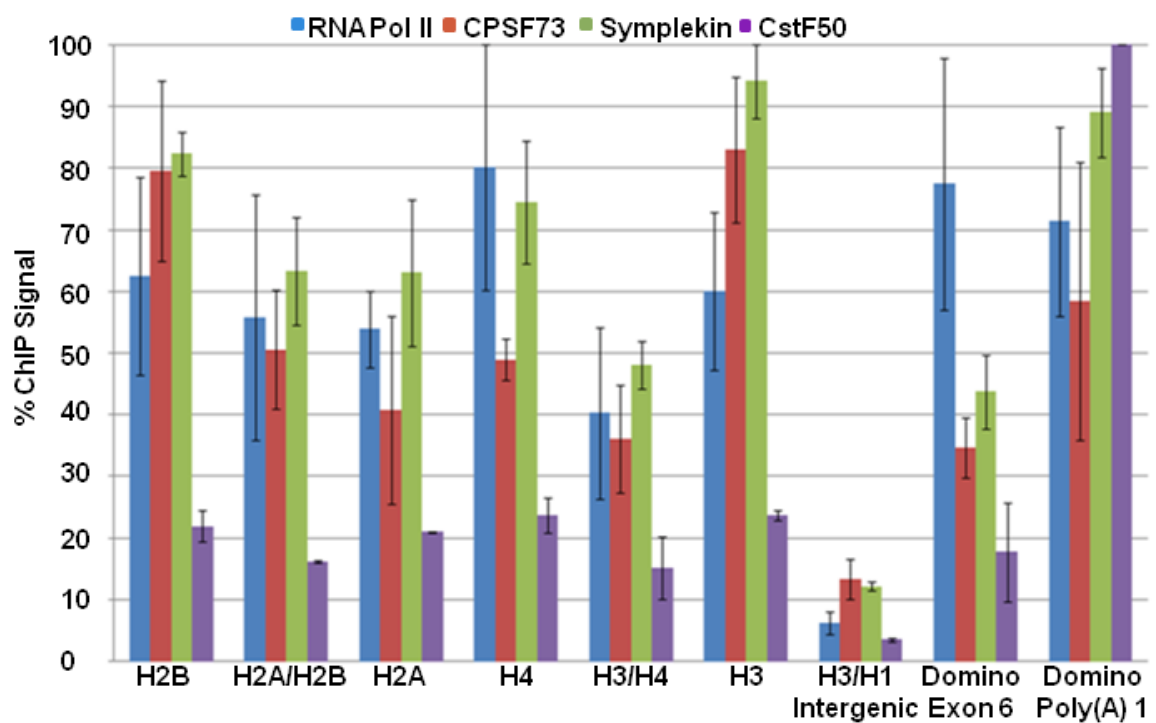
**Figure 12. RNA processing factors are recruited to histone genes in vivo. (A)**

Diagram of ChIP assay design. Oligonucleotide sets target short regions in the histone gene cluster. These sets correspond to the coding regions of H2A, H2B, H3 and H4, as well as the promoters of the H2A/H2B and H3/H4 pairs and the H3/H1 intergenic region. Control oligonucleotide sets were designed to exon 6 and the proximal poly(A) site of the domino gene. (B) ChIP assays were performed on crosslinked lysates from Dmel-2 cells using antibodies to the Rpb3 subunit of RNA polymerase II, CPSF73, Symplekin and CstF50. Columns represent average relative ChIP signal normalized to the maximal value for each antibody. Data are presented as the average of 3 independent experiments  $\pm$  SD.

**A**

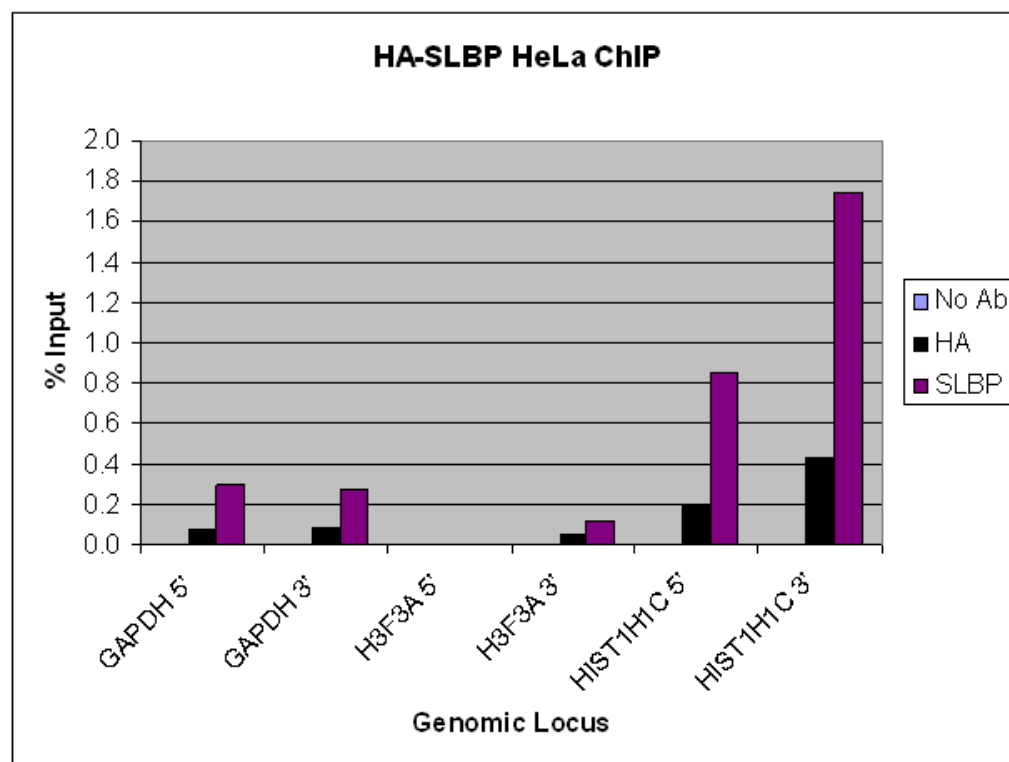


**B**

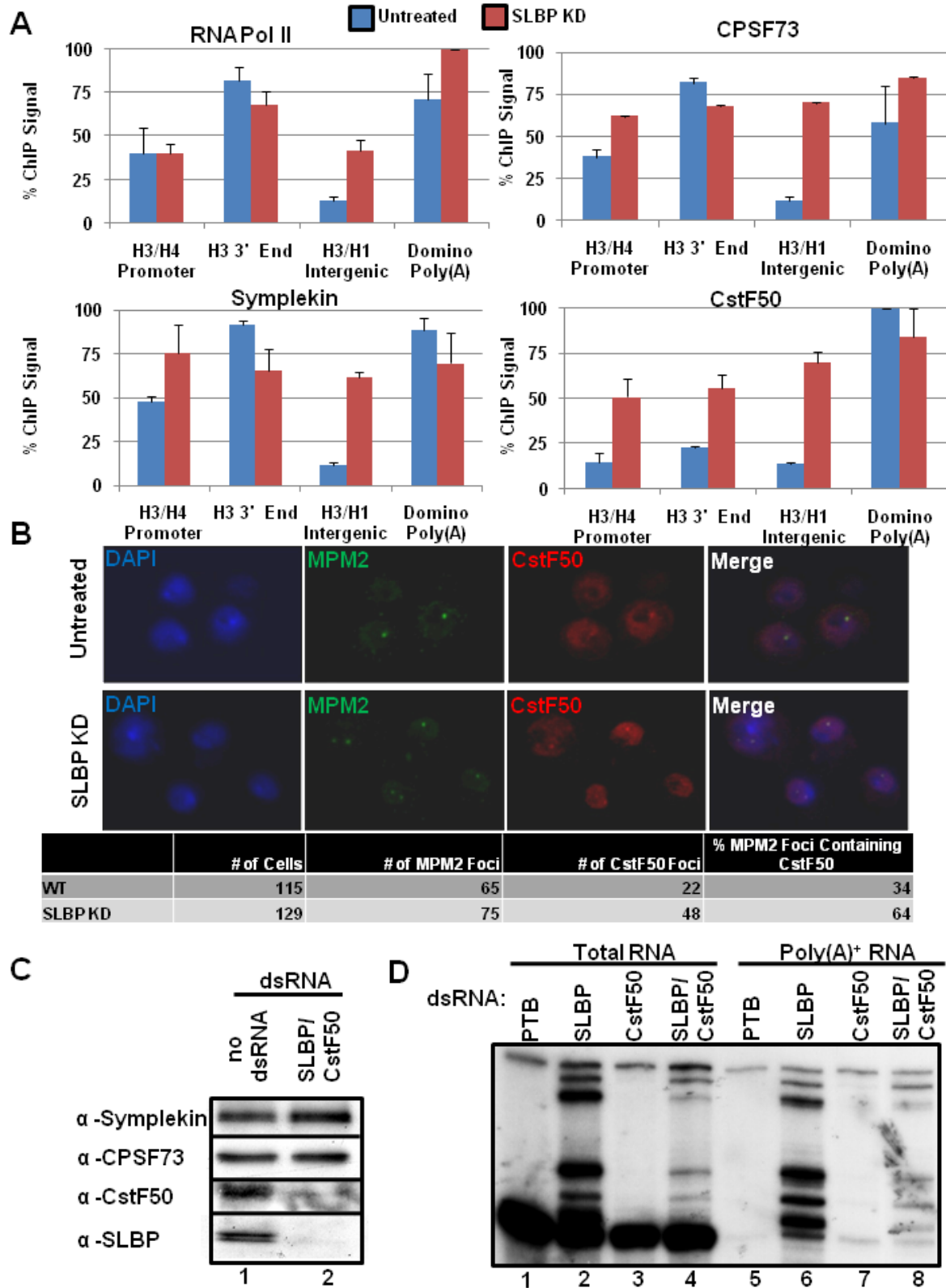


**Figure 13. SLBP associates with histone genes in mammalian cells.** ChIP

experiments were performed on HeLa lysates using HA (black bars) or SLBP (purple bars) antibodies.

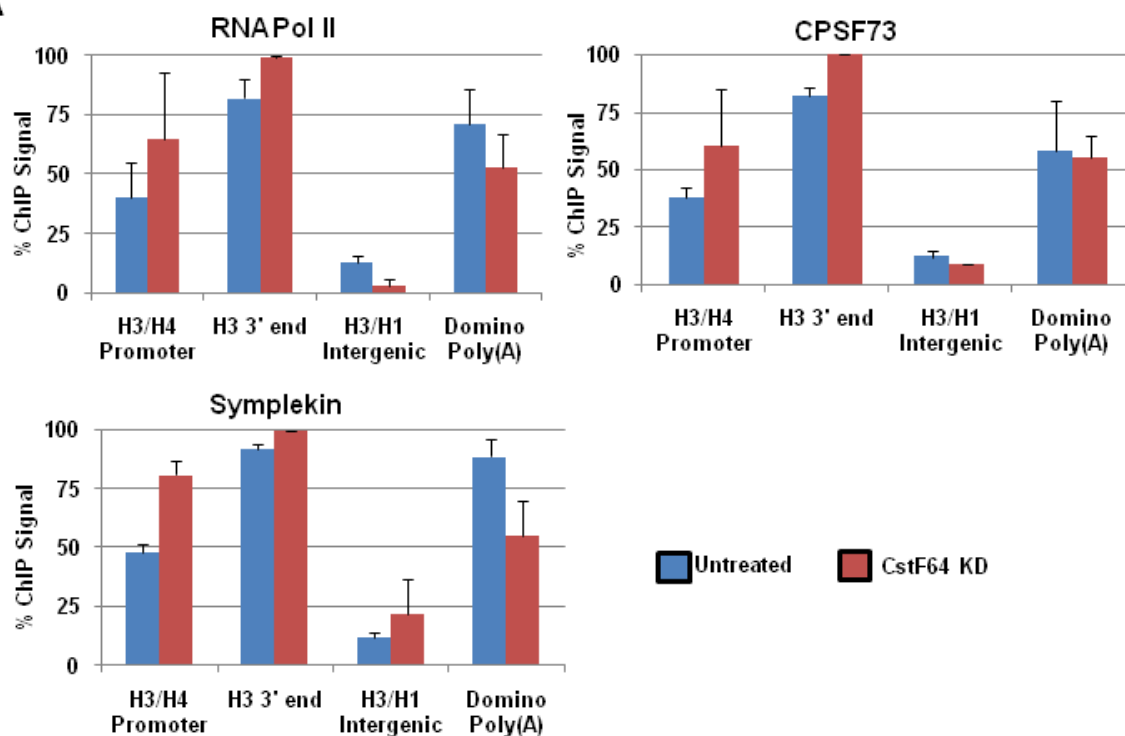


**Figure 14. Loss of a histone specific processing factor results in recruitment of additional poly(A) factors to histone genes.** (A) ChIP assays were performed using lysates prepared from Dmel-2 cells treated with control (blue) or SLBP (red) double-stranded RNA. Columns represent the average relative ChIP signal normalized to the maximal value for each antibody. Blue bars in each are untreated cells, red bars are values for SLBP knockdown cells. Data are presented as the average of 3 independent experiments and the SD was less than 20%. (B) Dmel-2 cells treated with control double-stranded RNA (Top) or SLBP double stranded RNA (Bottom) were fixed and probed for CstF50 (green) and Mpm2 (red), and also stained with DAPI (blue). The number of cells with Mpm2 foci, and the number of cells with Mpm2 foci that also contained CstF50 foci. There was a 2-fold increase in the fraction of cells containing Mpm2/CstF50 foci in the SLBP knockdown cells. (C) Dmel-2 cells were control treated with PTB dsRNA or dsRNA targeting SLBP and CstF50. Cell lysates were resolved by SDS-PAGE and western blots were performed. (D) Total RNA was prepared from cells treated with double-stranded RNAs against PTB, SLBP, CstF50 or both SLBP and CstF50. S1 nuclease protection assays were performed on total RNA (lanes 1-4) or poly(A)<sup>+</sup> RNA (lanes 5-8) as in Fig. 1C, D.

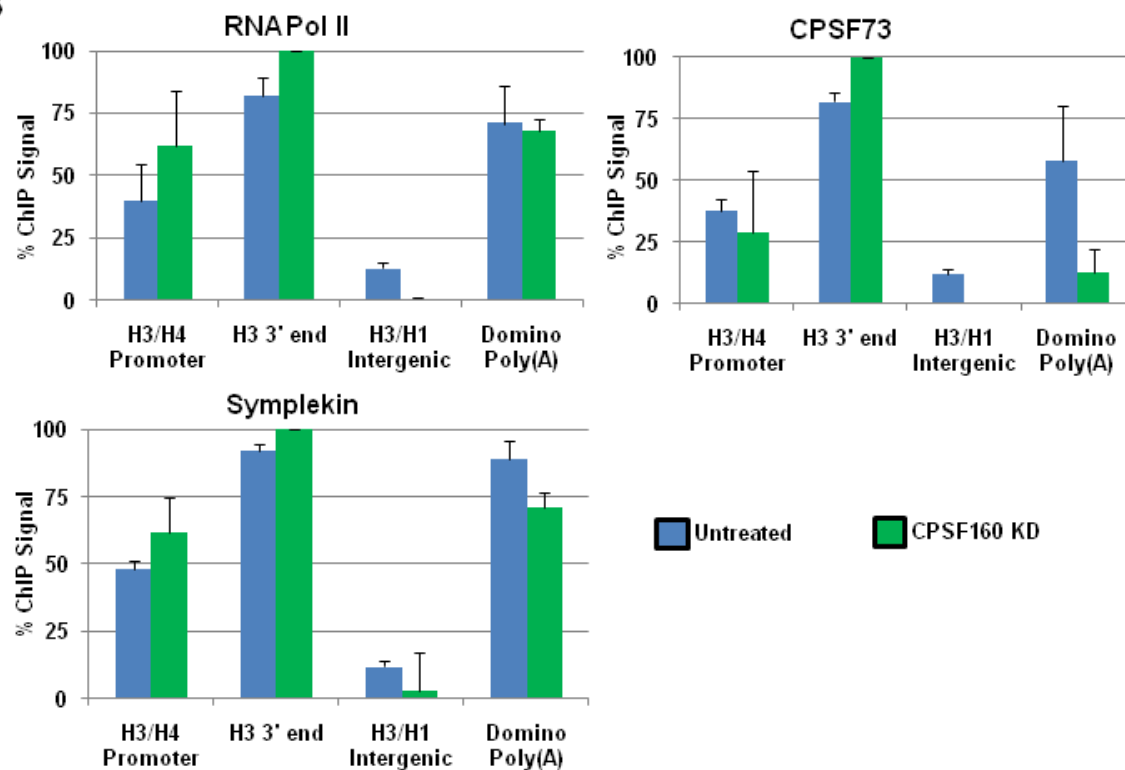


**Figure 15. Knockdown of poly(A) specific factors does not affect recruitment of core cleavage factor to histone genes.** Lysates from untreated cells (panel A and B, blue) and from cells treated with dsRNA targeting CstF64 (panel A, red) or CPSF160 (panel B, green) were made and ChIP assays were performed as in Fig. 3B. SD's were less than 20% among 3 different analyses.

**A**

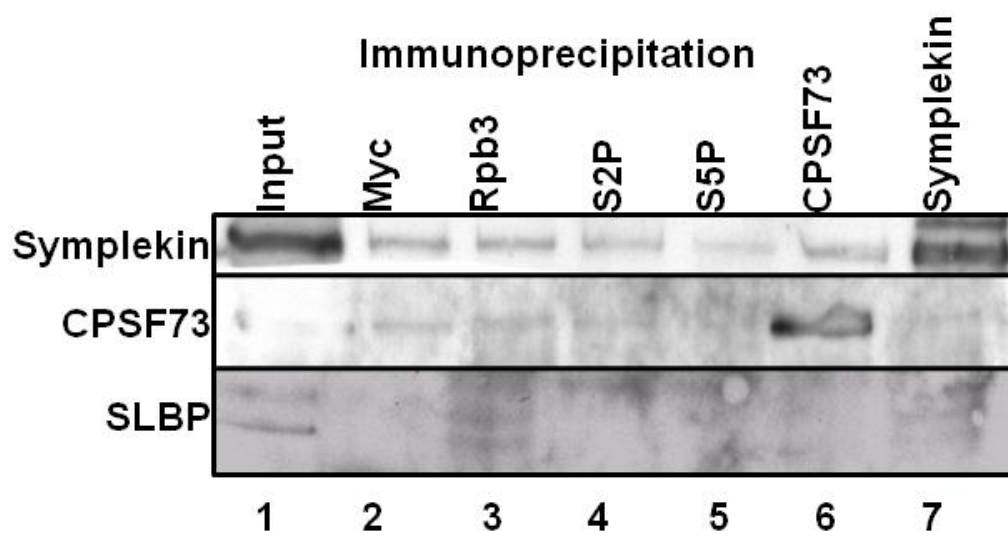
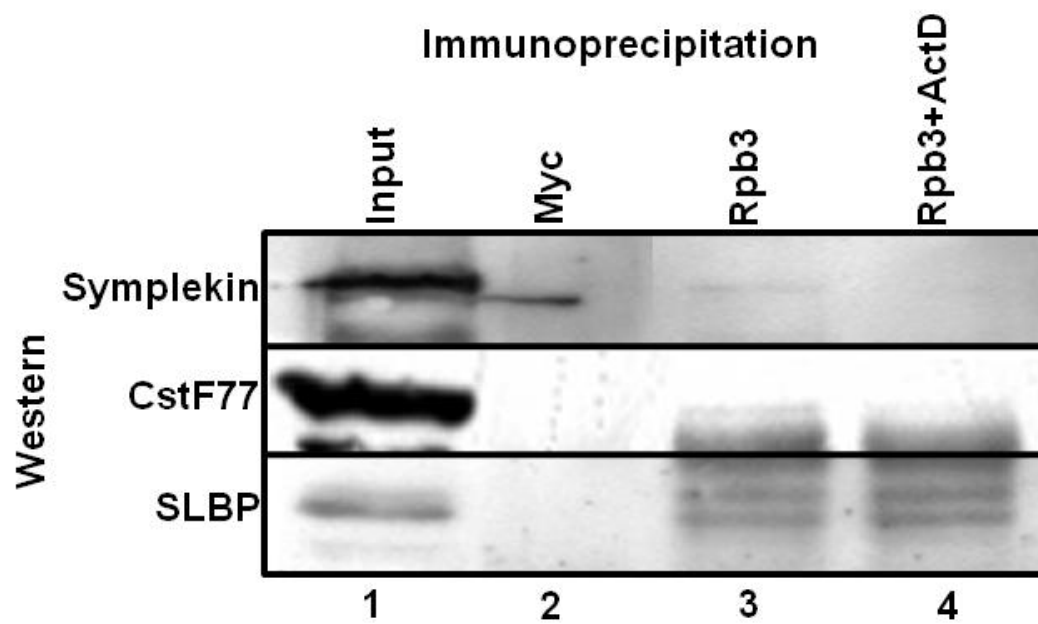


**B**





**Figure 17. 3' end processing factors associate with RNA Pol II.** (A) IPs were performed as in figure 9 using crosslinked whole cell lysates from Dmel-2 cells. Western blots were performed using the indicated antibodies. (B) IPs were performed as in figure 11 from untreated NEs or NEs from actinomycin D treated Dmel-2 cells.

**A****B**

**Figure 17. Model of histone pre-mRNA processing.** (A) Model of normal histone 3' end formation. (B) Additional poly(A) factors are recruited to histone genes to participate in cleavage and polyadenylation in the absence of histone specific RNA processing proteins.



## **Chapter 3**

### **CPSF73 is both the endonuclease and a processive 5' to 3' exonuclease in histone pre-mRNA processing**

#### **Introduction**

3' end processing is an essential step in converting eukaryotic mRNA precursors (pre-mRNAs) to mature mRNAs. Metazoan pre-mRNAs are processed at the 3' end by two distinct mechanisms (Zhao et al., 1999a). The vast majority of mRNA precursors are processed by cleavage coupled to polyadenylation. This mechanism depends on the presence of two sequence elements in the precursor; a highly conserved polyadenylation signal AAUAAA and a less conserved G/U-rich downstream element (Colgan and Manley, 1997). Cleavage occurs between the two sequence elements, usually about 20 nucleotides 3' to the AAUAAA. The upstream cleavage product is subsequently polyadenylated. Cleavage/polyadenylation is executed by a large complex composed of at least 12 proteins. The AAUAAA interacts with a 5-subunit cleavage/polyadenylation specificity factor (CPSF) (Kaufmann et al., 2004) whereas the G/U-rich downstream element interacts with a 3-subunit cleavage stimulation factor (CstF) (Colgan and Manley, 1997). Among other proteins, the cleavage/polyadenylation machinery also includes poly(A) polymerase that catalyzes

addition of the poly(A) tail (Colgan and Manley, 1997) and Symplekin, which plays a yet uncharacterized role in 3' end processing (Takagaki and Manley, 2000a). An increasing number of both genetic and biochemical studies indicate that endonucleolytic cleavage that proceeds polyadenylation is carried out by CPSF73, the 73 kDa subunit of CPSF (Mandel et al., 2006; Ryan et al., 2004). CPSF73 is a member of the metallo- $\beta$ -lactamase family of mostly hydrolytic enzymes that utilize zinc ions during catalysis (Aravind, 1999; Callebaut et al., 2002). A small group of the metallo- $\beta$ -lactamase proteins function in metabolism of DNA and RNA substrates as both endonucleases and/or 5'-3' exonucleases (Callebaut et al., 2002; Dominski and Marzluff, 2007; Mathy et al., 2007a).

A distinct 3' end processing mechanism operates on metazoan replication-dependent histone pre-mRNAs (Dominski and Marzluff, 2007). These transcripts are cleaved between a highly conserved stem-loop structure and a purine-rich histone downstream element (HDE) directly producing the mature histone mRNAs that end with the stem-loop followed by a 5-nucleotide single-stranded tail. 3' end processing of histone pre-mRNAs critically depends on the HDE, which interacts with U7 snRNP. The interaction is mediated by formation of a double stranded RNA between the HDE and the 5' end of U7 snRNA, an approximately 60-nucleotide component of the U7 snRNP. The U7 snRNP contains an unusual Sm complex, in which the Sm proteins D1 and D2 are replaced by two related proteins; Lsm10 and Lsm11 (Pillai et al., 2001; Pillai et al., 2003). The stem-loop structure forms a tight complex with the stem-loop binding protein (SLBP). SLBP also binds a 100 kDa zinc finger protein (ZFP100), which in turn interacts with Lsm11. This bridging interaction is believed to stabilize U7

snRNP on the HDE (Dominski and Marzluff, 2007). In common with cleavage and polyadenylation, 3' end processing of histone pre-mRNAs also involves Symplekin although its role remains to be elucidated (Kolev and Steitz, 2005).

UV-cross linking studies demonstrated that the cleavage site in histone pre-mRNAs interacts in a U7-dependent manner with CPSF73 thus strongly suggesting that the same endonuclease functions in generation of both polyadenylated and nonadenylated (histone) mRNAs (Dominski et al., 2005b). Following the in vitro cleavage reaction, the downstream product containing the HDE is rapidly degraded to mononucleotides by a nuclease that also depends on the U7 snRNA giving rise to mononucleotides (Dominski et al., 2005b). Our preliminary data suggested that also this reaction is catalyzed by CPSF73 (Dominski et al., 2005b).

In this study Zbigniew Dominski, Xiao Cui-Yang, and I show that degradation of the DCP is a result of a processive 5'-3' exonuclease activity and further support the involvement of CPSF73. We also identify important features of the RNA substrate that dictate whether CPSF73 functions as an endonuclease or 5' exonuclease and provide in vivo data suggesting that degradation of the DCP formed after histone pre-mRNA processing does not require by Xrn2, consistent with CPSF73 degrading the DCP.

## **Material and Methods**

**Preparation and labeling of RNA substrates.** Most RNAs were synthesized in Dharmacon (Colorado) and depending on the quality of synthesis, were used either directly or gel purified to isolate the full length product. RNA substrates containing a single internal label were constructed by ligating two RNA fragments using DNA ligase and a bridging oligonucleotide. Prior to ligation, the downstream fragment

was labeled at the 5' end with  $^{32}\text{P}$  using T4 polynucleotide kinase, as described (Dominski et al., 2005b). All RNA substrates were named based on the number of nucleotides present in front of the DCP, for example DCP, DCP+5 or DCP-2. Names of the substrates containing a single internal label additionally reflected the position of a modified nucleotide relative to the DCP sequence: 2'-O-methyl (m), deoxy (d) or phosphorothioate (s). The DCP+10 and DCP+15 substrates were constructed by ligating the 5' end of the DCP-9/s to a 19- or 24-nucleotide RNA fragment, respectively. The ligation products were gel purified, quantified and labeled at the 5' end with  $^{32}\text{P}$  using T4 polynucleotide kinase. The internally labeled 86-nucleotide H2a pre-mRNA containing the stem-loop structure and the HDE was synthesized by T7 transcription in the presence of  $^{32}\text{P}(\alpha\text{UTP})$ , as described (Dominski et al., 1999).

**Preparation of nuclear extracts and RNA processing and degradation .** Nuclear extracts were prepared from mouse myeloma cells, as described (Dominski et al., 1999). Unless otherwise indicated, all processing and degradation reactions were carried out 90 min in the presence of 20 mM EDTA, as described (Dominski et al., 1999). The efficiency of degradation was calculating by comparing the amount of the released radioactive mononucleotide to the input substrate used in the reaction. The U7-dependence of processing and degradation reactions was determined by using 100 ng of an anti-M oligonucleotide that is complementary to the first 17 nucleotides of the mammalian U7 snRNA (Dominski et al., 2005b). A 2'-O-methyl anti-D oligonucleotide directed to the first 19 nucleotides of the *Drosophila* U7 snRNA was used at the same concentration as a negative control (Dominski et al., 2005b). In some reactions, the



U7-dependence was also confirmed by using 500 ng (a 1,000-fold excess relative to the labeled substrate) of an RNA containing the wild type sequence of the HDE (wtHDE RNA). A corresponding RNA containing a 4-nucleotide mutation replacing the AAGA with UUCU within the HDE (mutHDE) was used at the same concentration as a negative control. Processing and degradation reactions were inhibited by NP-40 at a final concentration of 0.05%.

**Analysis of processing products.** After 90 min incubation, processing reactions were treated for 1 hour with proteinase K at 0.5 µg/µl, diluted 5 times with a 7M urea dye and analyzed in low resolution 8% polyacrylamide/7M urea gels to assess the reaction efficiency. The same samples were additionally analyzed in high resolution gels allowing for separation fragments that differ by one nucleotide.

Mononucleotides containing a phosphate at the 5' end were identified by electrophoresis in 12% gels based on co-migration with mononucleotides generated by S1 nuclease, which leaves a 5' phosphate, or products of KOH hydrolysis, which migrate 0.5 nucleotide faster due to the presence of an additional phosphate at the 3' end. The length of larger products was determined in high resolution 8% gels by comparison with appropriate size markers containing the same sequence (5' labeled RNAs used for ligation) or with products of partial KOH hydrolysis. In other cases, the length of final processing products was determined by comparison with the length of processing intermediates generated due to random stalling of the U7-dependent exonuclease at each nucleotide.

**Xrn1 ribonuclease assay.** The activity of yeast Xrn1 was tested in a buffer containing 30 mM Tris-HCl at pH 8.2, 2 mM MgCl<sub>2</sub>, 50 mM NH<sub>4</sub>Cl, 0.5 mM DTT, and 20 µg/ml

bovine serum albumin, as described (Mathy et al., 2007). Each reaction contained 1 ng of an RNA substrate and 10 ng of Xrn1 in a final volume of 10  $\mu$ L and was incubated at 32° C for 15 min.

**RNA interference.** RNAi was performed using a two-hit method (Wagner and Garcia-Blanco, 2002). Briefly,  $7 \times 10^6$  HeLa cells were plated into 24-well plates and transfected with 3  $\mu$ L of 20  $\mu$ M stock siRNAs (Xrn2-1, Xrn2-2, or a combination of both) 24 hours later. A day after the first siRNA treatment, cells were split 1:3 into 6-well plates and after another 24 hours were retransfected with 4  $\mu$ L of siRNA. 72 hours after the second transfection, samples were collected either by addition of Trizol (Invitrogen) or NP-40 lysis buffer directly to the wells. siRNA target sequences in the human Xrn2 mRNA were GGGAAGAAAUAUUGGCAA (Xrn2-1) and AAGAGUACAGAUGAUGAUG (Xrn2-2).

**qRT-PCR.** 2.5  $\mu$ g of total cell RNA was treated with DNase (Promega) and reverse transcription reactions were performed with MMLV-RT (Invitrogen) using random hexamers to prime the cDNA. cDNA from the reaction was added directly to qPCR reactions containing 2X SYBR GREEN PCR Master Mix (Applied Biosystems) and oligonucleotide primers. PCR was performed using a 7900HT PCR System (Applied Biosystems) and data were analyzed using SDS 3.2 software (Applied Biosystems).

The oligonucleotides used were: Hist2H2AA 3' End F:

GGAGCAGTACGGCCTGGAT, Hist2H2AA 3' End R:

CGACGAGGAACTGAACAAGCT, Hist2H2AA #1 Downstream F:

GGGACCCACTCATCGAAGAG, Hist2H2AA #1 Downstream R:

CGCGCCGTCTTCCATCT, Hist2H2AA #2 Downstream F:

CCAGGGCGCTTTGGAAA,

Hist2H2AA #2 Downstream R: GAGCTGGGTCTTGGCTTCAC, GAPDH 3' End F:

CCGCACCTTGTCATGTACCA, GAPDH 3' End R:

CCCTAGAATAAGACAGGACAAGTAACTG, GAPDH Downstream F:

TCGCTCCAGTCCTAGGCTATCT, GAPDH Downstream R:

GGCTGCCCACAGAATAGCTT.

## Results

### **The U7-dependent degradation of the DCP+1 substrate containing a site specific radioactive phosphate.**

To study 3' end processing of histone pre-mRNAs in vitro Zbig and Xiao use an 86-nucleotide derivative of the mouse H2a-614 pre-mRNA that contains all necessary processing elements, including the stem-loop and the HDE (Fig. 18A). The H2a substrate labeled by random incorporation with  $^{32}\text{P}$  UTP is processed in a mouse nuclear extract to an upstream cleavage product containing the stem-loop, and to 5' monophosphate nucleosides that result from the degradation of the downstream cleavage product (DCP) (Fig. 18B, lane 2)(Dominski et al., 2005b). 3' end processing of histone pre-mRNAs proceeds in the presence of 20 mM EDTA, which inhibits virtually all nonspecific nucleases present in the mouse nuclear extract. Blocking the 5' end of the U7 snRNA by a complementary 2'-O-methyl oligonucleotide (anti-M) at 10 ng/ $\mu\text{l}$  prevents binding of the U7 snRNP to the RNA substrate and abolishes processing (Fig. 18B, lane 3). A 2'-O-methyl oligonucleotide with an unrelated

sequence targeted to the 5' end of the *Drosophila* U7 snRNA (anti-D), routinely used at the same concentration to measure the extent of nonspecific effects, does not affect processing (not shown). A synthetic RNA substrate with a sequence corresponding to the DCP or extended at the 5' end by one or more nucleotides are also degraded in a U7-dependent manner in nuclear extracts from mammalian (Dominski et al., 2005b; Walther et al., 1998) and *Drosophila* cells (Dominski et al., 2005a). Thus, the degradation of the DCP can be uncoupled from the endonucleolytic cleavage of histone pre-mRNA.

In their previous studies Zbig and Xiao used a DCP+1 RNA labeled at the 5' end with a radioactive phosphate (Dominski et al., 2005a). The U7-dependent degradation of this substrate was monitored by the accumulation of  $^{32}\text{P}$  labeled mononucleotide. Since substrates labeled at the 5' end have a number of limitations, including the inability to follow degradation further into the body of the RNA, in the current studies Zbig and Xiao used a DCP+1 RNA containing a single internal radioactive phosphate. The 39-nucleotide DCP+1 RNA was constructed by ligating an 18-nucleotide upstream fragment (U. Half) with a 5' labeled 21-nucleotide downstream fragment (D. Half) (Fig. 18C). The DCP+1 RNA contained a hydroxyl group at the 5' end and a single radioactive phosphate between residues 17 and 18 of the DCP, near the center of the HDE. Incubation of this RNA in a mouse nuclear extract under standard conditions (20 mM EDTA) resulted in a release of the labeled mononucleotide from the body of the RNA with a small amount of the product visible after 5 min incubation (Fig. 18D, lanes 1-5). The nuclease functions processively since no intermediates were detected during the 60 min reaction. The anti-M oligonucleotide at 10 ng/ $\mu\text{l}$  blocked degradation of the

DCP+1 substrate whereas the anti-D oligonucleotide at the same concentration had no effect (Fig. 18D, lanes 6-8) demonstrating that nuclease activity depends specifically on binding of the U7 snRNP to the HDE.

In order to release the radioactive nucleotide, the nuclease had to remove nearly half of the HDE that is required to recruit the U7 snRNP. Zbig and Xiao carried out a competition experiment to determine whether RNAs containing only partial HDE sequences can bind the U7 snRNP and therefore inhibit processing of the H2a-614 pre-mRNA. Zbig and Xiao used the downstream and the upstream DCP halves (D. Half and U. Half, respectively) (Fig. 18C), each at a 1,000-molar excess over the pre-mRNA substrate. Neither the D. Half nor the U. Half RNAs, containing 9 and 6 nucleotides of the HDE, respectively affected the processing of the pre-mRNA substrate (Fig. 18E, lanes 5 and 6). In contrast, the same molar excess of the DCP+1 RNA containing the entire HDE prevented cleavage of the H2a-614 pre-mRNA (Fig. 18E, lane 4). Thus, the removal of the first six nucleotides of the HDE during the degradation of the DCP likely displaces the U7 snRNP from the RNA substrate. Zbig and Xiao also tested whether 5' labeled D. Half and U. Half RNAs can be degraded in a mouse nuclear extract in a U7-dependent manner. Consistent with the inability of partial HDE sequences to bind the U7 snRNP, neither substrate generated U7-dependent degradation products (not shown).

In Fig. 17C the DCP+1 substrate was labeled internally with a single radioactive phosphate and contained a hydroxyl group at the 5' end. The fact that this substrate was efficiently degraded demonstrates that the U7-dependent nuclease does not require a phosphate at the 5' end. To determine whether the addition of a 5' phosphate

can stimulate the degradation efficiency, Zbig and Xiao incubated the DCP+1 RNA with T4 polynucleotide kinase and unlabeled ATP and tested the resulting 5' phosphorylated substrate in a nuclear extract. The efficiency of degradation was not significantly increased indicating that the U7-dependent nuclease activity does not prefer substrates with a 5' phosphate.

**The exonuclease degrades a substrate containing a radioactive phosphate downstream of the HDE.**

The above experiments demonstrated that the U7-dependent 5' exonuclease can remove at least the first 19 nucleotides of the DCP+1 RNA releasing labeled 5' UMP. To determine whether the exonuclease can progress further into the body of the RNA substrate to remove the entire U7 binding site Zbig and Xiao used a 60-nucleotide substrate (3'extDCP) created by ligating the DCP+1 RNA to a 5' labeled 21-nucleotide RNA designated 3'ext. In the 3'extDCP substrate the only radioactive phosphate was located 12 nucleotides downstream of the 3' end of the HDE (Fig. 19A). Incubation of the 3'extDCP substrate in a mouse nuclear extract resulted in a U7-dependent accumulation of labeled 5' AMP indicating that the 5' exonuclease can remove at least 40 nucleotides from the 5' end (Fig. 19B, lane 6), progressing completely through the HDE and displacing the U7 snRNP. Compared to the degradation of the DCP+1/m<sub>19</sub> (Fig. 19B, lane 2), the degradation of the 3'extDCP was less efficient. Since the two substrates have the same sequence for the first 39 nucleotides, including the HDE, the most likely explanation for the inefficient degradation of the 3'extDCP is an inhibitory effect of the additional 21 nucleotides located at the 3' end on the recruitment of the U7 snRNP (see also below). As observed for other substrates, the degradation efficiency

of the 3'extDCP substrate can be significantly improved by the presence of a high excess of the mutHDE RNA (Fig. 19B, lane 12).

### **Switching between the 5' exonuclease and endonuclease modes.**

In Zbig and Xiao previous study we used a DCP substrate that was extended at the 5' end by 5 nucleotides and contained a phosphorothioate modification at the cleavage site (Dominski et al., 2005b). This substrate, previously called HDE+5/s and renamed here to DCP+5/s<sub>cs</sub> (Fig. 20A), when labeled at the 5' end and incubated in a mouse nuclear extract yielded labeled mononucleotide diagnostic of the 5' exonuclease (Dominski et al., 2005b). There was no detectable 5-nucleotide fragment produced demonstrating that this substrate is not cleaved endonucleolytically (Dominski et al., 2005b). In the nuclear extract used in these studies, incubation of the DCP+5/s<sub>cs</sub> substrate again predominantly yielded labeled mononucleotide (Fig. 20B, lane 1). One explanation for the predominant 5' exonucleolytic degradation was that the presence of phosphorothioate at the cleavage site makes endonucleolytic attack by CPSF73 less favorable than cleavage at the most proximal bond lacking any modification. To test this possibility Zbig and Xiao designed a new substrate, DCP+5/s<sub>5'</sub>, containing a phosphorothioate modification at the first phosphodiester bond rather than within the cleavage site (Fig. 20A). However, also this substrate was nearly exclusively cleaved after the first nucleotide releasing labeled mononucleotide. These results indicate that CPSF73 preferentially utilizes the 5' exonucleolytic mode of its activity if a substrate contains only five nucleotides upstream of the cleavage site. To determine whether the predominant cleavage after the first nucleotide results from the inherent inability of the DCP+5 substrate to undergo endonucleolytic cleavage, Zbig and Xiao next analyzed

the pattern of degradation of DCP+5/m<sub>5</sub> RNA in which the first nucleotide was modified by a 2'-O-methyl group (Fig. 20A). This modification renders phosphodiester bonds fully resistant to hydrolysis by CPSF73 and hence it was expected to block the 5' exonuclease activity. The U7-dependent products generated by incubation of this substrate in a nuclear extract were resolved by high resolution gel electrophoresis next to products of complete digestion of the same RNA by yeast 5' exoribonuclease Xrn1 and endonuclease S1 (Fig. 20B, lanes 3 and 4). These two nucleases generate a phosphate at the 5' end and thus their products serve as appropriate size markers for the CPSF73 generated products. To unambiguously determine the size of the degradation products Zbig and Xiao also used a mixture of two chemically synthesized fragments: mACC and mACCC (Fig. 20B, lane 6). As expected, degradation of the DCP+5/m<sub>5</sub> RNA in the nuclear extract did not release labeled mononucleotide and instead generated a group of major U7-dependent products consisting of 2, 3 and 4 nucleotides (Fig. 20B, lane 2). There was also a small amount of a 5-nucleotide product (arrow) that resulted from hydrolysis at the natural endonucleolytic cleavage site. Interestingly, as judged by the accumulation of labeled mononucleotide, the first phosphodiester bond in the DCP+5/m<sub>5</sub> RNA was readily hydrolyzed by Xrn1 (Fig. 20B, lane 3). This could be a biologically relevant feature of Xrn1 5' exonuclease as many mRNAs that are degraded by Xrn1 in vivo contain 2' O-methyl modified nucleotides near the 5' end as a part of the cap structure. We conclude that CPSF73 strongly prefers to cleave after the first nucleotide of the DCP+5 RNA although it is also inherently capable of cleaving this substrate endonucleolytically.



The minimal H2a pre-mRNA substrate that we routinely use to analyze 3' end processing in vitro contains 38 nucleotides upstream of the cleavage site, including the stem-loop (Fig. 18A). To determine the minimal length upstream of the cleavage site that allows endonucleolytic cleavage, Zbig and Xiao constructed DCP+10 (Fig. 20A) and DCP+15 (Fig. 20C) RNA substrates. In these substrates, the cleavage site was preceded by 10 or 15 nucleotides found in the same position in the H2a pre-mRNA (Fig. 18A). Incubation of the DCP+10 substrate with a mouse nuclear extract yielded a U7-dependent product that migrated significantly slower than mononucleotides (Fig. 20C, lane 3). Generation of each product was blocked by the anti-M oligonucleotide thus it was dependent on the U7 snRNP (Fig. 20C, lane 4). Electrophoresis in a high resolution gel next to appropriate size markers revealed that DCP+10 was cleaved endonucleolytically within the expected site generating a 10-nucleotide product (Fig. 20D, lane 5). The DCP+15 RNA was also cleaved endonucleolytically generating a 15 nucleotides long product {I would put this in the paper (Fig. 20D, lane 3).

Thus cleavage is endonucleolytic if the 5' end of the transcript extends 21 nt of the start of the HDE, and exonucleolytic if the 5' end of the transcript extends between 9 and 16 nts 5' of the HDE. Consistent with previously published results (Mowry et al., 1989), no secondary structure is required for endonucleolytic cleavage. The cleavage site is the same as found in the histone pre-mRNA, consistent with the previous demonstration (Scharl and Steitz, 1994) that the position of the HDE (and hence the U7 snRNP) determines the cleavage site.

**Xrn2 knockdown stabilizes downstream products in GAPDH but not histone pre-mRNAs in vivo.**

According to the “torpedo model” for termination of transcription by RNA polymerase II on polyadenylated mRNAs, degradation of the downstream cleavage product in the nascent transcript promotes release of the RNA polymerase II from the DNA template resulting in termination of transcription (Buratowski, 2005b; Luo and Bentley, 2004; Rosonina et al., 2006; Tollervey, 2004). In support of this model, the nuclear 5' to 3' exonuclease Xrn2 has been implicated in transcription termination in yeast and mammalian cells (Kaneko et al., 2007; Kim et al., 2004; Luo et al., 2006b; West et al., 2004). Our in vitro results suggest that CPSF 73 rather than Xrn2 could be the 5' exonuclease that degrades the downstream cleavage product in histone pre-mRNAs in vivo potentially promoting termination of transcription. To test this possibility, I used two different siRNAs targeted to the Xrn2 mRNA and efficiently reduced the level of Xrn2 in HeLa cells (Fig. 21A, lanes 2-4). I analyzed the effect of knockdown of Xrn2 on the downstream cleavage products of both the polyadenylated GAPDH transcript and the Hist2H2AA transcripts. Each of these genes is quite close to another gene and hence transcription termination must occur within ~1000 nucleotides of the 3' cleavage site (Fig. 21A). The transcription termination site of the mouse orthologue of the Hist2H2AA gene has been mapped about 600 nucleotides after the cleavage site (Chodchoy et al., 1991). I designed oligonucleotides that target a region of the GAPDH pre-mRNA 233-295 nucleotides downstream from the cleavage site as well as primers sets which target regions 140-197 and 457-526 nucleotides downstream of the Hist2H2AA cleavage site (Fig. 21B). To determine the relative amounts of the downstream cleavage product to the mature mRNA I also used

oligonucleotides which amplify target sequences upstream of the 3' cleavage site in each transcript (3' end regions).

To determine what effect Xrn2 knockdown had on stabilization of downstream products of GAPDH and Hist2H2AA pre-mRNAs, qRT-PCR was performed.

Relative amounts of downstream products were determined by comparing  $C_T$  values of downstream PCR reactions with those of the 3' end reactions for each knockdown and normalizing these to the values from control treated cells. Typically, I observed a difference in  $C_T$  values of about 8 cycles between the downstream fragment and the 3' fragment from the mRNA in control cells, implying the concentration of the downstream fragment was between 0.1 and 1% mature mRNA, and the levels of the mature mRNA did not change in the Xrn2 knockdown cells. This analysis allowed me to determine  $\Delta\Delta C_T$  values which represent relative differences in the amounts of downstream products in the knockdowns compared to control cells. Negative values denote fewer cycles for the PCR reaction to reach the threshold value in knockdown cells than control cells, indicating that was a higher concentration of the RNA species in the sample. I used two different siRNAs against Xrn2 as well as a mixture of the two, and they all reduced with Xrn2 levels, with the combination being most effective (Fig. 21B, lane 2). Xrn2 knockdown had little effect (less than 1 cycle difference) on the amounts of Hist2H2AA downstream products (Fig. 21C). In contrast, there was a substantial increase in the amount of downstream product from the GAPDH transcript (indicative of its stabilization) in all three knockdowns of Xrn2. Knockdown of Xrn2 with either individual siRNA resulted in a stabilization of ~2 cycles while the combination knockdown revealed a change of 4 cycles (Fig. 21C),

consistent with the degree of knockdown of Xrn2 observed by Western blotting. Taken together, these data indicate the Xrn2 is involved in the degradation of the GAPDH downstream product but not the Hist2H2AA downstream products. This is consistent with the possibility that the 5' exonuclease activity of CPSF73 (and not Xrn2) is involved in degradation of the downstream product after cleaving the pre-mRNA. Note that knockdown of any component of the 3' end processing machinery would result in the increase of the downstream fragment due to failure to process the pre-mRNA, and hence direct analysis of the effect of knocking down of CPSF73 is not possible.

Thus our results strongly support the idea that CPSF73 is responsible for endonucleolytic cleavage, and then processive degradation of the downstream fragment assisting in the termination of transcription by RNA polymerase II.

## **Discussion**

### **The U7-dependent nuclease degrades the downstream cleavage product in the 5' to 3' direction and stalls at modified nucleotides.**

Histone pre-mRNAs are cleaved at the 3' end by a multi-component processing machinery that includes the U7 snRNP bound to a region downstream of the cleavage site (Dominski and Marzluff, 2007). Following cleavage in vitro, the downstream cleavage product (DCP) is rapidly degraded to mononucleotides by a nuclease that also depends on the U7 snRNP. To demonstrate that degradation is carried out by a 5' to 3' exoribonuclease Zbig and Xiao incorporated differently modified nucleotides into specific sites in the DCP RNA: 2'-O-methyl nucleotides, deoxynucleotides and

phosphorothioate nucleotides. Each modification resulted in the accumulation of intermediate products whose size clearly demonstrated that degradation begins from the 5' end and stalls at the modified nucleotide, thus it behaves as a typical 5' exonuclease. The strongest block was created by a 2'-O-methyl modification, which nearly completely stopped progression of the exonuclease. A single deoxynucleotide located in the same position was relatively well tolerated by the U7-dependent exonuclease whereas the phosphorothioate modification had an intermediate ability to block the degradation of the DCP.

The U7-dependent 5' exonuclease activity is processive and proceeds past the HDE, suggesting that it can readily displace the U7 snRNP from the substrate. Thus, binding of the U7 snRNP to the HDE is only required to recruit the 5' exonuclease and initiate degradation. It is not known whether the exonuclease acts by itself to displace the U7 snRNP or requires assistance from other proteins present in the nuclear extract. However, given that the reaction occurs in the presence of EDTA, the involvement of a conventional helicase in unwinding the RNA duplex formed between the U7 snRNA and the HDE seems unlikely.

**The distance of the 5' end of the transcript from the HDE determines whether CPSF73 acts as an endonuclease or an exonuclease.**

UV-cross-linking studies strongly suggest that the downstream cleavage product is degraded by CPSF73. Cross-linking of this protein to the first scissile bond in the DCP RNA was dependent on the U7 snRNP, but also perfectly correlated with the efficiency of degradation. CPSF-73 was also cross-linked to the

DCP-2/s RNA that contains only 9 nucleotides upstream of the HDE and corresponds to a substrate partially truncated from the 5' end.

Initiation of exonucleolytic degradation occurs when the 5' end of the substrate is between 9 and 16 nts 5' of the start of the HDE. A substrate that contains 21 nts 5' of the HDE is cleaved endonucleolytically (with the downstream cleavage product subsequently degraded exonucleolytically). Surprisingly, a DCP derivative containing 16 nucleotides 5' of the start of the HDE (DCP+5), and which contains the normal endonucleolytic cleavage site, ACCCA, was also degraded from the 5' end. Endonucleolytic cleavage of this substrate could be activated by blocking degradation from the 5' end of the DCP+5 RNA with a 2'O-methyl group. Thus, CPSF73 preferentially uses the 5' exonuclease mode on the DCP+5 RNA although it is intrinsically capable of cleaving this substrate endonucleolytically. Interestingly, as many as four major sites located close to the 5' end were cleaved upon blocking the 5' end of the DCP+5, with the natural cleavage site being used only rarely. Thus, the processing complex assembled on the DCP+5 substrate leaves CPSF73 flexibility in selecting sites for endonucleolytic cleavage. Importantly, exclusive cleavage at the correct site was achieved by adding 5 more nucleotides to the 5' end of the DCP+5 RNA (DCP+10 substrate) instead of blocking this end by a 2'O-methyl group. Perhaps, the region upstream of the cleavage site in the DCP+10 substrate is sufficiently long to interact with a component of the processing machinery, or with a site on CPSF73 itself, which may stabilize its interaction with CPSF73 directing cleavage to the correct site. An important role in preventing the 5' exonuclease mode in processing of the DCP+10 RNA may also be played by the fact that the 5'

end is located too far away from the HDE and thus not easily accessible for CPSF73. Similarly, if the number of nts 5' of the start of the HDE is 7 or less, the substrate is not degraded, suggesting that CPSF73 cannot bind these substrates and initiate degradation (although it readily degrades these substrates during processive, exonucleolytic degradation of the DCP. Thus failure to degrade these substrates is due to the inability to initiate degradation.

Two critical observations argue against the possibility that the two known and closely related 5' exonucleases, Xrn1 and Xrn2, are directly responsible for the in vitro degradation of the DCP in histone pre-mRNA. First, both Xrn1 and Xrn2 depend on divalent ions and are completely inactivated by low concentrations of EDTA (Solinger et al., 1999; Stevens and Poole, 1995). Second, the 5' exonuclease activity of Xrn1 or Xrn2 requires the presence of a 5'phosphate (Solinger et al., 1999; Stevens, 1978; Stevens, 1980; Stevens and Poole, 1995). We confirmed that these two features indeed characterize Xrn1 from *S. cerevisiae* (a gift from L. Benard and C. Condon) using a set of our RNA substrates (data not shown). In contrast to Xrn1 and Xrn2 nucleases, the U7-dependent 5' exonuclease efficiently degrades substrates containing a 5' hydroxyl.

While the U7-dependent activity degrading the DCP in histone pre-mRNA differs from the Xrn 5' exonucleases, it is similar to RNase J1, which together with CPSF73 belong to the  $\beta$ -CASP subfamily of the metallo- $\beta$ -lactamase proteins (Callebaut et al., 2002; Mandel et al., 2006). RNase J1 was initially characterized as an endonuclease involved in processing of mRNA precursors in *Bacillus subtilis* (Even et al., 2005) and subsequently shown to also function as a 5' to 3' exonuclease in mRNA degradation

and rRNA processing (Mathy et al., 2007b). The ability to act as both an endonuclease and a 5' to 3' exonuclease is also shared by DNA-specific nucleases of the  $\beta$ -CASP subfamily; Artemis, Apollo and Snm1 (Dominski, 2007). Altogether, this analysis strongly supports the notion that CPSF73 possesses both endonuclease and 5' exonuclease activities.

In contrast to our results, a bacterially expressed N-terminal fragment of human CPSF73 functions only as an endonuclease and fails to degrade RNA substrates in the 5'-3' direction (Mandel et al., 2006). One explanation is that the 5' exonuclease activity of CPSF73 requires an assistance of a second component of the processing complex and can be only observed in nuclear extracts or reconstituted systems. It is also possible that the 5' exonuclease (but not the endonuclease) activity of CPSF73 vitally depends on the C-terminal portion of the protein. This interpretation is supported by recent crystallographic studies, which demonstrated the importance of the C-terminal domain for the activity of RNase J1 (de la Sierra-Gallay IL et al., 2008).

Previously, an activity that degrades the DCP during histone pre-mRNA processing has been identified in mammalian nuclear extracts by Walther et al (Walther et al., 1998). In contrast to our studies, this activity stalled at guanine nucleotides and hence generated a family of fragments trimmed at the 5' end rather than mononucleotides. Further studies are required to determine its relation to CPSF73 although the dependence on the U7 snRNP and resistance to EDTA suggest that the two activities represent the same enzyme.

**Xrn2 is not required for degradation of the DCP in histone pre-mRNA in vivo.**



Recent in vivo and in vitro studies demonstrated that the degradation of the downstream cleavage product during cleavage/polyadenylation involves the well characterized 5' exonuclease, Xrn2 (Kaneko et al., 2007; Kim et al., 2004; Luo et al., 2006a; West et al., 2004). I demonstrated that downregulation of Xrn2 by RNAi does not result in significant stabilization of the downstream cleavage product generated from the Hist2H2AA transcript while as expected the corresponding product from the transcript encoding GAPDH was in the same cells was stabilized by knocking down Xrn2, in agreement with previous studies of polyadenylated mRNAs (Luo et al., 2006a; West et al., 2004). Thus, it is likely that the 5' exonuclease activity of CPSF73 degrades the DCP in histone pre-mRNA also in vivo while the same function for all other protein-encoding transcripts is carried out by Xrn2. The difference may reflect the fact that 3' end processing of conventional pre-mRNAs is controlled by the upstream AAUAAA signal, which recruits CPSF subunits, including CPSF73 (Murthy and Manley, 1995), whereas in 3' end processing of histone pre-mRNA the recruitment of CPSF73 critically depends on the downstream HDE, which interacts with the U7 snRNP. Following processing of histone pre-mRNA CPSF73 likely remains associated with the U7 snRNP and is perfectly positioned to use its preferential 5' exonuclease activity to degrade the downstream cleavage product. In polyadenylated mRNAs, the CPSF73 remains associated with the rest of CPSF and the upstream product while it is undergoing polyadenylation. It is possible that Xrn2 plays a supportive role in the degradation process by acting on histone transcripts after random dissociation of CPSF73. This interpretation is consistent with our in vivo results demonstrating that the most distal region of the histone transcript is

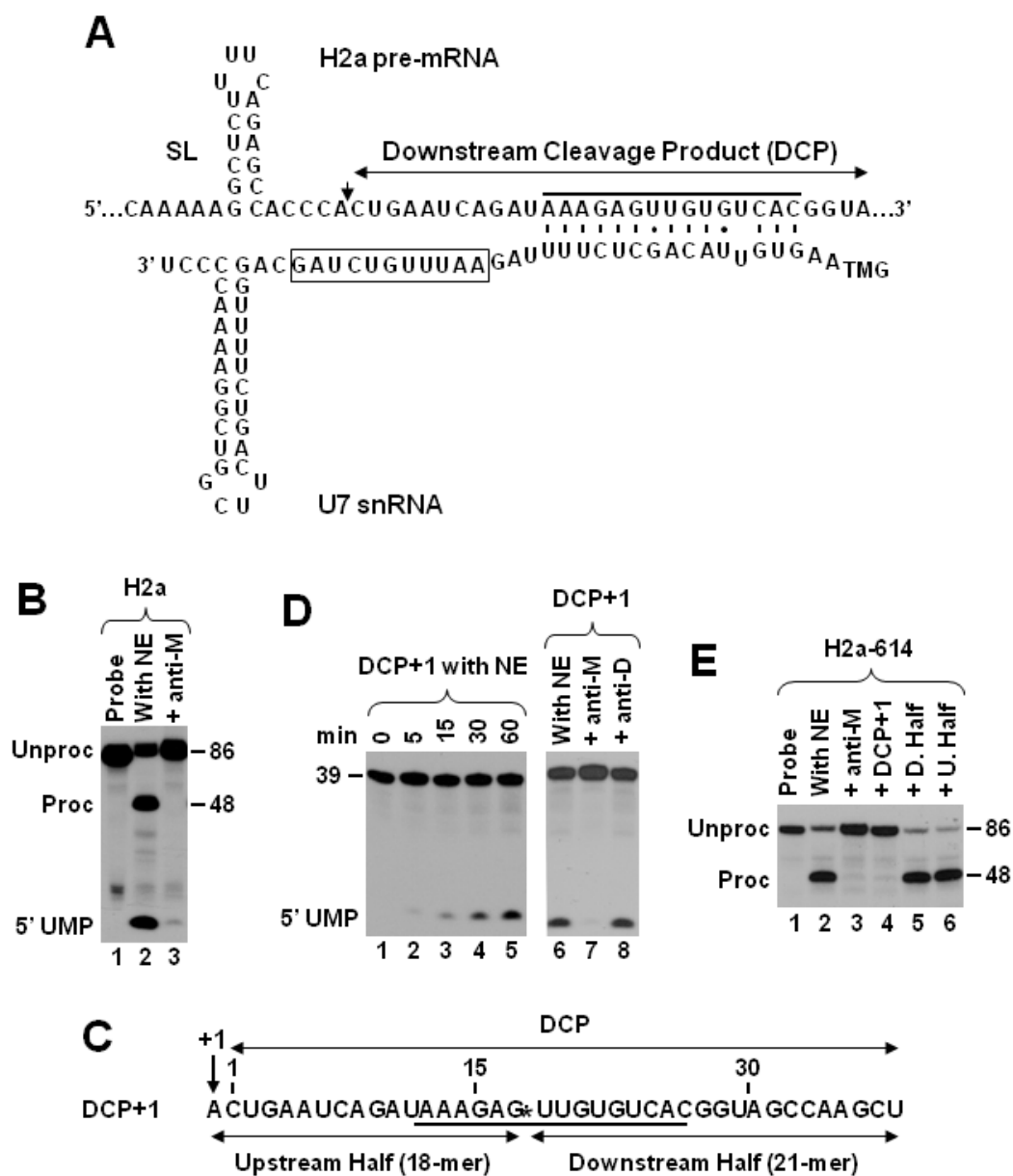
moderately stabilized by siRNA-mediated depletion of Xrn2, although the levels of the region 200 nts from the cleavage site is not altered by Xrn2 (Fig. 21C).

One obvious function for the 5' exonuclease activity of CPSF73 *in vivo* may be to degrade the region immediately downstream of the cleavage site and therefore to liberate the U7 snRNP from its base pairing association with the HDE for another round of 3' end processing, as previously suggested (Walther et al., 1998).

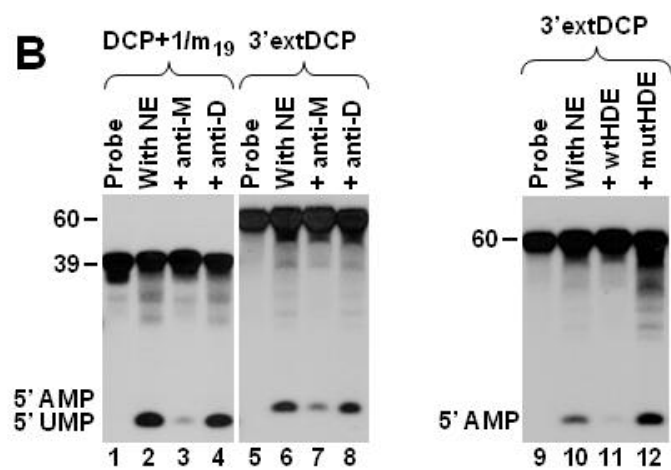
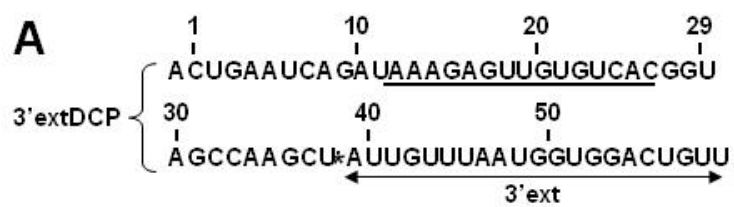
Moreover, an attractive possibility is that the CPSF73 continues degradation into more downstream regions of the nascent histone pre-mRNA ultimately resulting in the displacement of RNA polymerase II from the DNA template and termination of transcription, as postulated for Xrn2 in the torpedo mechanism. Indeed, mutations within the 3' end processing elements in histone pre-mRNAs, the stem-loop and the HDE, prevent transcription termination *in vivo* (Chodchoy et al., 1991), suggesting that as in cleavage/polyadenylation (Buratowski, 2005b) 3' end processing of histone pre-mRNA is functionally linked to the release of RNA polymerase II from histone genes.

**Figure 18. The U7-dependent degradation of the downstream cleavage product generated during 3' end processing of histone pre-mRNAs.** **A.** A fragment of the 86-nucleotide mouse histone H2a pre-mRNA substrate encompassing processing elements (top sequence) and the base pair interaction between the HDE (indicated with a thick line) and the U7 snRNA (bottom sequence). The cleavage site and the downstream cleavage product (DCP) are indicated with an arrowhead and a double arrowed line, respectively. The Sm site in the U7 snRNA is indicated with a box. **B.** In vitro processing of the internally labeled 86-nucleotide H2a substrate in a mouse nuclear extract (NE) under control conditions (lane 2) or in the presence of an oligonucleotide blocking the U7 snRNA (anti-M, lane 3). Lane 1 contains the unprocessed input pre-mRNA (Unproc). The 48-nucleotide processing product (Proc) results from cleavage of the input after the fifth nucleotide 3' of the stem-loop. **C.** The sequence of the DCP+1 RNA containing a single radioactive phosphate between nucleotides 17 and 18 of the DCP sequence (see Materials and Methods for the terminology used to designate RNA substrates). The DCP+1 RNA was constructed by ligating the upstream and the downstream halves, with the downstream half containing the radioactive phosphate at the 5' end (indicated with an asterisk). The HDE interacting with the U7 snRNA is underlined. **D.** A 60-min time course of the U7-dependent degradation of the DCP+1 RNA in a mouse nuclear extract (lanes 1-5) and the effect of an oligonucleotide blocking either the mouse U7 snRNA (anti-M, lane 7) or *Drosophila* U7 snRNA, used as a nonspecific control (anti-D, lane 8). **E.** In vitro processing of the 5' labeled H2a-614 pre-mRNA either under control conditions (lane 2) or in the presence of a 1,000-fold

molar excess of RNA competitors containing either complete (lane 4) or partial HDE (lanes 5-6). Processing shown in lane 3 was carried out in the presence of the anti-M oligonucleotide.



**Figure 19. Degradation of the 3'extDCP RNA containing a single radioactive phosphate downstream of the HDE. A.** The sequence of the 60-nucleotide 3'extDCP RNA. The HDE is underlined and the position of the radioactive phosphate is indicated with an asterisk. The 3'ext RNA is indicated with a double arrowed line. **B.** The U7-dependent degradation of the 3'extDCP RNA under control conditions (lane 6 and 10) or in the presence of various RNA competitors, as indicated at the top of each lane. The degradation efficiency for the DCP+1/m<sub>19</sub> RNA (Fig. 3A) is shown in lanes 1-4 for comparison.



**Figure 20. Switching between the 5' exonuclease and endonuclease modes. A.**

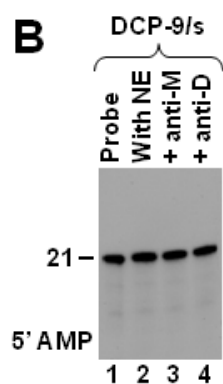
The sequence of DCP derivatives extended at the 5' end by 5 or more nucleotides and labeled at the 5' end (asterisk). The position of phosphothioate (s) and 2'O-methyl (m) modifications are indicated. **B.** A high resolution gel analysis of the U7-dependent products generated by incubation of the DCP+5/s<sub>CS</sub> (lane 1) and DCP+5/m<sub>5'</sub> (lanes 2 and 5) substrates in a mouse nuclear extract. Lanes 3 and 4 contain products of complete digestion of the DCP+5/m<sub>5'</sub> RNA with the yeast Xrn1 5' exonuclease or S1 nuclease, respectively. Lane 6 contains a mixture of two synthetic oligonucleotides labeled at the 5' end, mACCC and mACCC that serve as size markers. **C.** The U7-dependent cleavage of the DCP+10 substrate under control conditions or in the presence of the anti-U7 oligonucleotides, as indicated. Degradation of the DCP-2/s (lane 1) substrate is shown to indicate the position of mononucleotides.



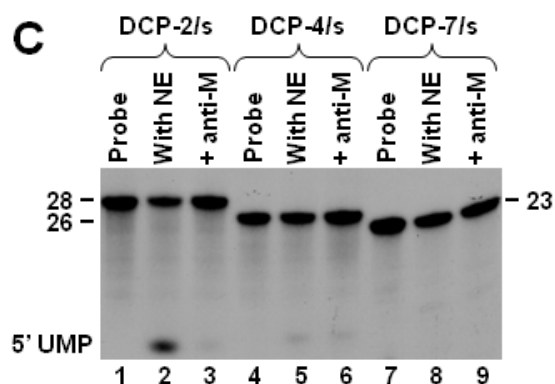
**A**

DCP-9/s \*A<sub>s</sub>UAAAGAGUUGUGUCACGGUA  
 DCP-7/s \*A<sub>s</sub>GAUAAAGAGUUGUGUCACGGUA  
 DCP-4/s \*A<sub>s</sub>UCA GAUAAAGAGUUGUGUCACGGUA  
 DCP-2/s \*U<sub>s</sub>AAUCA GAUAAAGAGUUGUGUCA CGGUA  
 DCP-2max/s \*U<sub>s</sub>AAUCA GAUAAAGAGcUGUGaCA CGGUA

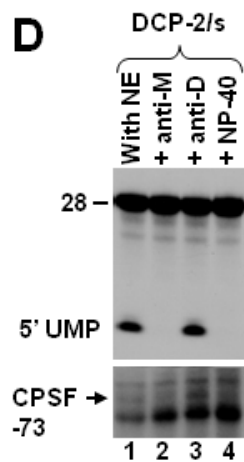
**B**



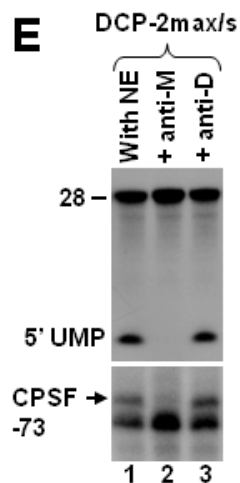
**C**



**D**



**E**



## Figure 21

### **Xrn2 is not required for degradation of Hist2H2AA downstream cleavage**

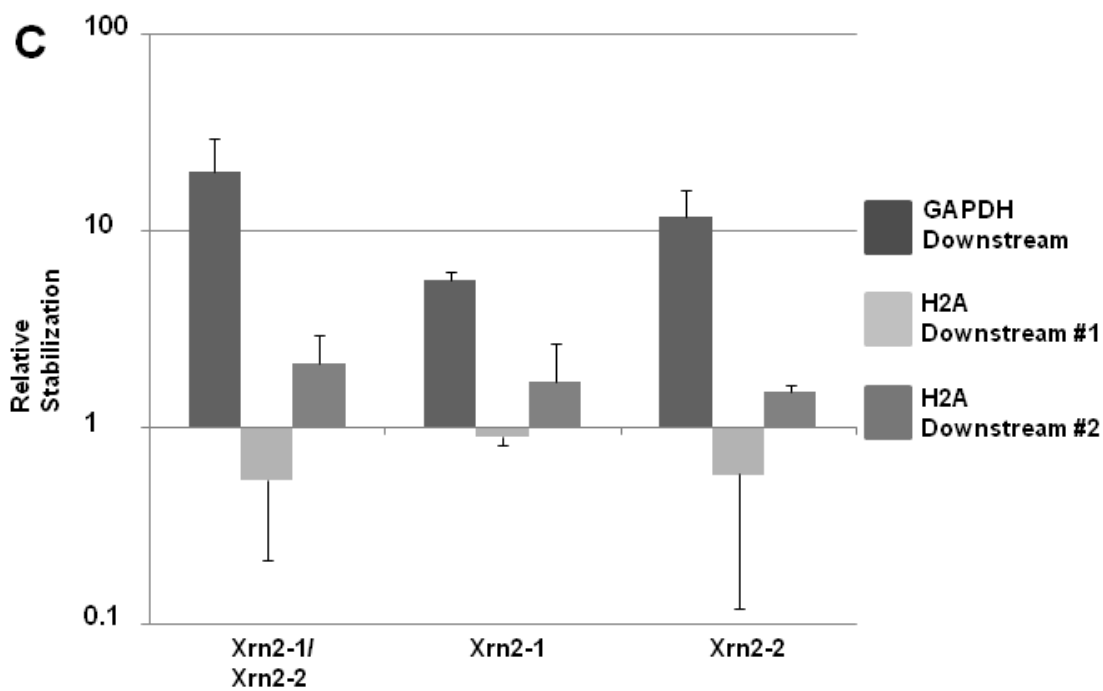
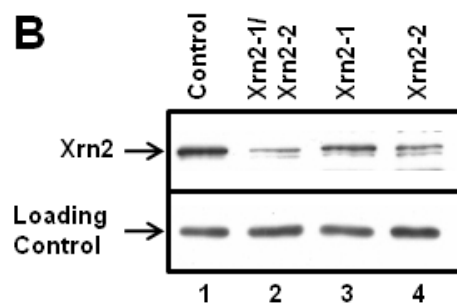
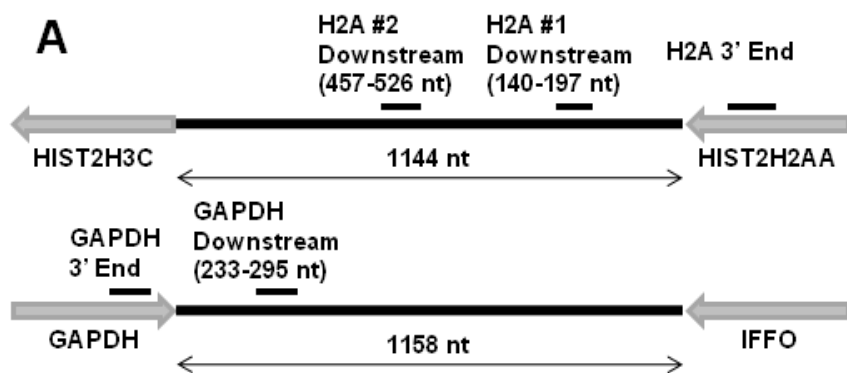
**product in vivo. A.** siRNA-mediated downregulation of Xrn2 in HeLa cells.

Individual Xrn2-1 and Xrn2-2 siRNAs (lanes 3 and 4) as well as their combination (lane 2) were used to reduce levels of Xrn2. HeLa cells were collected 72 hours after second siRNA transfection and subjected to Western blot analysis using an  $\alpha$ -Xrn2 antibody. A cross-reacting band detected with the Xrn2 antibody serves as the loading control.

**B.** Probe design for qRT-PCR. Primers were designed to probe regions within the mature mRNAs (3' End) of both GAPDH and Hist2HAA for normalization purposes. Additional oligonucleotide sets were designed targeting regions 140-197 nucleotides and 457-526 nucleotides downstream of Hist2H2AA cleavage site (H2A #1 Downstream and H2A #2 Downstream, respectively) as well as a region 233-295 nucleotides downstream of the GAPDH cleavage site (GAPDH downstream).

**C.** HeLa cells were treated as in (A) and RNA prepared by Trizol extraction. RT reactions were performed using MMLV-RT and resultant cDNAs were subjected to qPCR using primers described in (B).  $\Delta\Delta C_T$  values were calculated using the equation  $(C_{T\text{siRNA}_{\text{Downstream}}} - C_{T\text{siRNA}_{3'\text{End}}}) - (C_{T\text{Control}_{\text{Downstream}}} - C_{T\text{Control}_{3'\text{End}}})$  where  $C_{T\text{siRNA}_{\text{Downstream}}}$  and  $C_{T\text{siRNA}_{3'\text{End}}}$  represent  $C_T$  values for a given siRNA treated sample amplified with the indicated primer set and  $C_{T\text{Control}_{\text{Downstream}}}$  and  $C_{T\text{Control}_{3'\text{End}}}$  represent  $C_T$  values for control treated samples amplified using the same indicated primer set. The absolute  $C_T$  value for the fragment within both the GAPDH and histone mRNAs varied by less than one cycle among the different samples, indicating there was not a significant change in

either mature mRNA as a result of the siRNA treatments. A negative  $\Delta\Delta C_T$  value indicates the treatment resulted in an increase in the relative amount of the RNA fragment. The results are an average of 3 independent experiments.



## **Chapter 4**

### **SLBP Knockdown Results in Nuclear Retention of Histone mRNA**

#### **INTRODUCTION**

The replication-dependent histone mRNAs are the only eukaryotic cellular mRNAs that are not polyadenylated, ending instead in a conserved stemloop (Marzluff, 2005). These mRNAs lack introns and undergo a single processing reaction cleaving the nascent pre-mRNA from the DNA template, and producing the mature mRNA. The cleavage reaction requires the stem loop (SL) 5' of the cleavage site and a purine-rich region, the histone downstream element (HDE), located about 10 nts 3' of the cleavage site. The HDE basepairs with the U7 small nuclear RNA, which is a component of the U7snRNP (Mowry and Steitz, 1987). The stem loop binding protein (SLBP) binds to the SL and helps stabilize the interaction of U7 snRNP with the pre-mRNA. Several lines of evidence support this model. Pulldown of biotinylated histone pre-mRNA from nuclear extracts results in the concomitant pulldown of the U7snRNA (Dominski et al., 1999; Dominski et al., 2001) which is enhanced by SLBP. In vitro processing of histone pre-mRNA is absolutely dependent on SLBP if the HDE does not interact strongly with U7snRNA

(Dominski et al., 1999). The factor ZFP100 interacts with both the SLBP/SL complex (Dominski et al., 2002) as well as the U7snRNP specific Sm-like protein Lsm11 (Azzouz et al., 2005; Wagner et al., 2006) and is limiting for histone pre-RNA processing in vivo (Wagner and Marzluff, 2006). Mutations in the stemloop that abolish binding of SLBP result in no expression of processed histone mRNA in vivo (Pandey et al., 1994).

The requirement of SLBP for in vitro processing of histone pre-mRNA in nuclear extracts from mammalian cells on many substrates is not absolute. Extracts depleted of SLBP are incapable of processing the H1t histone pre-mRNA but are competent to process the H2A-614 histone pre-mRNA (Dominski et al., 1999). The differences between these two substrates is that U7snRNP basepairs more extensively with the HDE of H2A-614 than with the H1t HDE.

Once the U7snRNP is bound to the HDE, a cleavage factor is recruited containing at least CPSF73, (Dominski et al., 2005b) and Symplekin (Kolev and Steitz, 2005). CPSF73 is likely the endonuclease that cleaves the histone pre-mRNA from the DNA template (Dominski et al., 2005b). The other components of the cleavage factor and the mechanism of recruitment of the cleavage factor to the U7snRNP/SLBP/SL complex have not been elucidated.

The molecular events that take place following processing in the nucleus and prior to the translation of histone mRNA in the cytoplasm are less well understood. What is clear is that the histone mRNA is rapidly exported (within 5 min; (Schochetman and Perry, 1972)) and export in *Drosophila* S2 cells (Erkman et al.,

2005a) and *Xenopus* oocytes (Huang and Steitz, 2001) is dependent on TAP. SLBP is present in the cytoplasm only when histone mRNA is present (Erkman et al., 2005b), where it is specifically associated with the histone mRNP (Townley-Tilson et al., 2006; Whitfield et al., 2004). In *Xenopus* oocytes, expression of SLBP stimulates the translation of reporter genes ending in a stem loop (Sanchez and Marzluff, 2002). Recently it has been reported that SLBP interacts with a novel protein, SLIP1, which also interacts with EIF4G and plays a role in translation of histone mRNA (Cakmakci et al., 2008). These data are all consistent with a role of SLBP in the translation of histone mRNA.

In late S phase or when DNA replication is inhibited, histone mRNAs are rapidly degraded. The rapid decay of existing histone mRNA likely requires SLBP. Upf1 binds to SLBP at the 3' end of histone mRNA immediately after inhibition of DNA replication and is required for histone mRNA degradation (Kaygun and Marzluff, 2005a). Recently we reported that an oligo(U) tail is added to the 3' end of histone mRNA shortly after inhibiting DNA replication (Mullen and Marzluff, 2008). This oligo(U) tail provides a binding site for Lsm1-7 which binds to the oligouridylated histone mRNA together with SLBP. The Lsm1-7 complex likely recruits the enzymes that degrade histone mRNA as it does for polyA mRNAs (Parker and Song, 2004). SLBP is a cell-cycle regulated protein which accumulates just before entry into S-phase to activate histone pre-mRNA processing and SLBP is degraded at the end of S-phase. Stabilization of SLBP by mutating the phosphorylation sites required for degradation does not affect degradation of histone

mRNA (Zheng et al., 2003), indicating that SLBP degradation is not required for histone mRNA degradation.

In order to understand the critical roles of SLBP in cells, Tom Mullen, Eric Wagner and I knocked down SLBP using RNA interference in U2OS cells.

Previously, we and others had shown that reduction of SLBP levels results in a reduction in the rate of cell growth and an accumulation of cells in S-phase (Wagner et al., 2005; Zhao et al., 2004). However, a thorough molecular analysis of the affect of SLBP knockdown on histone mRNA metabolism has not been reported. Here we show the major lesion in histone mRNA metabolism as a result of knocking down SLBP is not in histone pre-mRNA processing, but in the export of histone mRNA to the cytoplasm. Properly processed histone mRNA accumulates in the nucleus. Together these results implicate SLBP as having a critical histone mRNA export function in mammalian cells, as well as its known roles in histone pre-mRNA processing, histone mRNA translation and regulation of histone mRNA half-life.

## **MATERIALS AND METHODS**

### **RNA interference**

RNAi was performed using a two-hit method as previously described using Lipofectamine 2000 (Wagner and Garcia-Blanco, 2002). Briefly, 110,000 U2OS or 95,000 HeLa cells were plated into 24 well plates and the following day, siRNAs were transfected into cells. A day following the first hit of siRNA, cells were replated into 6 well dishes, allowed to grow an additional day and then transfected with the second hit of siRNA. Cells were harvested 48-72 hours following the second hit for



analysis. The siRNA sequences of the C2 (control) and S2 (SLBP) have been previously described (Wagner et al., 2005).

### **DNA Synthesis and Mitotic Index Measurements by FACS**

We quantified the percentage of bromodeoxyuridine (BrdU) labeled cells in S phase and phospho-histone H3 in mitosis as previously described (Doherty et al., 2003). As controls for these assays cells were treated with 5 mM HU for 1 hour prior to addition of the BrdU. To measure DNA synthesis in S phase, RNAi-treated cells were pulsed with a 1 mM stock solution of BrdU (10  $\mu$ M final concentration) for 1 hour. Acid-treated nuclei were resuspended with 100  $\mu$ l of a solution containing  $\alpha$ -BrdU antibody conjugated to FITC (Becton Dickinson) diluted 1:5 in IFA (10 mM HEPES pH 7.4, 150 mM NaCl, 4% FBS, 0.1% NaN<sub>3</sub>) plus 0.5% Tween-20 and incubated in the dark, at room temperature for 2 hours. Following the incubation, 2 ml of IFA plus 0.5% Tween-20 was added and nuclei recovered by centrifugation. Nuclei were resuspended in 500  $\mu$ l of the IFA plus 0.5% Tween solution containing 5  $\mu$ g/ml RNase A and 50  $\mu$ g/ml propidium iodide. Nuclei were incubated overnight at 4°C and analyzed on a Becton-Dickinson (Becton Dickinson) FACScan using Summit Software (Dako Cytomation).

To measure the mitotic index using phospho-histone H3 as a marker, cells fixed with 1% formaldehyde in PBS for 30 minutes. Fixed cells were pelleted at 1200 rpm for 5 minutes at 4°C and the supernatant removed. Cells were resuspended in 1.5 ml PBS and then 3 ml of cold 100% ethanol was added and cells stored at 4°C overnight. The cells were resuspended in 100  $\mu$ l of a solution

containing 0.5  $\mu\text{g}$  of phospho-histone H3 (serine 10) antibody (Upstate Biotechnology) and 5  $\mu\text{g}/\text{ml}$  RNase A in IFA plus 0.5% Tween-20 for 2 hours at room temperature. The cell pellet was resuspended in 100  $\mu\text{l}$  of the IFA plus 0.5% Tween-20 solution containing an anti-rabbit secondary antibody conjugated to FITC (Santa Cruz Biotechnology) for 1 hour at room temperature. The cells were resuspended in 500  $\mu\text{l}$  of the IFA plus 0.5% Tween solution containing 5  $\mu\text{g}/\text{ml}$  RNase A and 50  $\mu\text{g}/\text{ml}$  propidium iodide and analyzed by FACS.

### **Northern Blot Analysis and S1 Nuclease Protection Assays**

Northern blot analysis to replication-dependent histone genes and 7SK snRNA was performed as described previously (Mullen and Marzluff, 2008). Briefly, 2  $\mu\text{g}$  of total RNA was resolved on 6% urea-polyacrylamide gels, transferred to Hybond N+ nitrocellulose (Amersham Biosciences), and probed in QuickHyb (Stratagene) overnight at 60°C. Probes were generated by random primed,  $\alpha^{32}\text{P}$ -dCTP labeling (Primelt II Kit, Stratagene) of PCR products generated to the complete ORF of cloned human histone genes and to 7SK RNA (Mullen and Marzluff, 2008).

S1 nuclease protection assays to map the 3' ends of histone H2a (HIST2H2AA) and H3 (HIST2H3A/C) were performed as previously described (Kaygun and Marzluff, 2005a). Probes for 3' end labeling were generated by digestion of the HIST2H2AA gene with *Ascl* and the HISTH3A/C gene with *BssHI*. The digested plasmid DNA was 3' end labeled with  $\alpha^{32}\text{P}$ -dCTP and Klenow (New

England BioLabs) and hybridized with total cell RNA. Protected fragments were resolved on 6% urea-polyacrylamide gels.

### **Western Blot Analysis**

Protein lysates from RNAi treated cells were prepared as previously described (Mullen and Marzluff, 2008) and the concentrations quantified using Bradford reagent. Primary antibodies used in this study include rabbit anti-SLBP (Wang et al., 1996), mouse anti-Symplekin (BD Biosciences), rabbit anti-fibrillarin (Abcam), rabbit anti-Hsp90 (Cell Signaling Technology), rabbit anti-histone H3 (Abcam). Protein were resolved on 12% SDS-polyacrylamide gels and transferred to nitrocellulose before detection using appropriate primary and secondary antibodies.

### **In Situ Hybridization**

In situ hybridizations were performed as described (Erkmann et al., 2005a). RNAi treated cells were grown on coverslips and washed with PBS and fixed in 3.7% PFA in PBS for 10 min. at room temperature. Cells were then washed again and permeabilized in 0.5% Triton X-100 in PBS. Pre-hybridization was performed for 10 min. at 37°C using a solution of 4x SSC and 50% formamide preheated to 37°C. Cover slips were inverted onto a 5-10 µl drop of hybridization solution (50% deionized formamide, 10% dextran sulfate, 1x Denhardt's solution, 4x SSC, 1 mg/ml E. coli tRNA) containing 1 ng/µl DIG labeled H2A or H3 probe. Probes were denatured for 5 min. at 80°C and cooled on ice prior to addition to hybridization solution. Hybridizations were performed overnight at 37°C in a humidified chamber.

Cells were washed once with 2X SSC at 37 °C, 3 times in 60% formamide in 0.2x SSC at 37°C, once [??] in 2 X SSC at room temperature, and 3X with 0.5% Tween-20 in 4X SSC. Coverslips were inverted onto 8 µl drops of 0.5% Tween-20 in 4X SSC containing mouse Cy3 α-DIG antibody diluted 1:200 and incubated at 37°C for 1h in a humidified chamber. Coverslips were washed 3X with 0.5% Tween-20 in 4X SSC and inverted onto 8 µl drops of with 0.5% Tween-20 in 4X SSC containing Cy-3 α-mouse secondary antibody diluted 1:200. Secondary incubation was performed for 45 minutes at 37°C in a humidified chamber. Coverslips were washed 3 X with 0.5% Tween-20 in 4X SSC before being fixed with 3.7% PFA in PBS for 5 min. at room temperature. After a PBS wash, coverslips were DAPI stained and mounted onto slides using Fluorsave reagent (Calbiochem). Images were acquired using a Zeiss LSM 510 confocal microscope and analyzed using ImageJ software.

### **CAT Assay**

CAT assays were performed essentially as described (Cullen, 2004).  $3 \times 10^5$  293T cells were co-transfected with 25 ng of reporter plasmid and 500 ng of MS2-fusion plasmid using Lipofectamine-2000. 48h after transfection cells were harvested in TEN buffer (100 mM NaCl, 10 mM EDTA, 10 mM Tris-HCl, pH 7.5), pelleted and resuspended in 100 µL of 100 mM Tris-HCl, pH7.5). Cells were lysed by freeze-thawing three times and lysates were pelleted for 5 min at 4°C. Equal amounts of lysates were added to 300 µL CAT assay buffer (5 mg/mL chloramphenicol in 100 mM Tris-HCl, pH 7.5 containing 1 µL/mL [<sup>3</sup>H]acetyl coenzyme A). 3 mL Econofluor (Packard) scintillation cocktail was added to each sample and counts determined in a scintillation counter.

## RESULTS

### SLBP is required for cell growth

I determined the effect of knocking down SLBP in U2OS cells using a double-stranded siRNA that specifically targets SLBP mRNA. Previously we have shown that knockdown of SLBP with this siRNA results in an increase in S-phase cells, and that we can rescue this phenotype by expressing an RNAi-resistant SLBP (Wagner et al., 2005). Moreover, Muller and colleagues (Zhao et al., 2004) demonstrated that reduced expression of SLBP also resulted in a defect in S phase progression and a reduced expression of histone mRNAs. Both studies demonstrated S phase defects but did not elucidate the specific molecular defects underlying this phenotype.

Since we initially observed variable results, which were dependent on the extent of SLBP knockdown, we optimized the knockdown, both for extent and duration allowing us to assay for histone mRNA metabolic defects, while still maintaining efficient SLBP knockdown. Eric and I used Western blotting to monitor SLBP protein levels over various times throughout the RNAi transfection (Fig. 22A). Reproducible phenotypes required a knockdown of SLBP >95%, as determined by comparing Western blot intensities from SLBP knockdown cells and a series of dilutions of the control lysates (Fig. 22B). U2OS cells were plated in 24 well dishes and the following day, transfected with either control (C2) or SLBP (S2) siRNA. We determined the levels of SLBP every 24 hrs. The day after the first transfection (24 hours, Figure 22A, lanes 2 and 3) SLBP protein levels were slightly lower in C2 treated cells than in untransfected cells (Figure 22A, lane 2). This decrease is likely due to the high cell density used to achieve effective transfection, resulting in a

reduction in the number of S-phase cells which contain SLBP. After 24 hours we split the cells out of the 24 well dishes and replated them into 6-well tissue culture dishes. 24 hrs later (48 hours after first siRNA treatment), we transfected the second dose of siRNA and monitored SLBP levels for the next 4 days. Between 72 and 120 hours following the initial hit of siRNA, SLBP protein levels were downregulated >95%, and experiments assaying histone mRNA metabolism were done during this time. Note that there is clearly some residual SLBP present in the cells at all times. By 144 hrs, SLBP levels had started to recover (data not shown).

As previously shown, knockdown of SLBP to this level resulted in an increase in the number of cells in S-phase (Wagner et al., 2005). I determined the rate of growth of the cells by measuring cell number at the various time points in Figure 22A. The SLBP knockdown cells grew more slowly, resulting in 30-40% fewer cells after 5 days of growth. All the cells were viable, and there was no evidence of apoptosis, suggesting that the cells were growing at a slower rate.

### **Reducing the levels of SLBP decreases the rate of DNA synthesis and activates an S-phase cell cycle checkpoint.**

To confirm that the SLBP knockdown cells were accumulating in S-phase Tom, Eric and I analyzed these cells by flow cytometry. We also asked whether these cells were still synthesizing DNA by pulsing the C2 or S2 siRNA treated U2OS cells with bromodeoxyuridine (BrdU), a thymidine analog, for 30 minutes followed by analysis of incorporation by FACS, using a FITC coupled anti-BrdU antibody. As a control we treated cells transfected with C2 (control) siRNA with 5 mM hydroxyurea (HU) for 1 hour to abolish BrdU incorporation. The DNA content was measured by

propidium iodide staining, and the incorporation of BrdU was measured by FACS. As demonstrated previously, the SLBP knockdown cells accumulated in S-phase of the cell cycle (Figure 22B, top panel). From the FACS analysis of BrdU-incorporation (Figure 22B, bottom panel), it is clear that SLBP knockdown cells are still synthesizing DNA (all the S-phase cells incorporated BrdU), but they were synthesizing DNA at about  $\frac{1}{2}$  the rate as control cells, as judged by the intensity of BrdU staining. In the control RNAi (C2) cells, there is efficient incorporation of BrdU as detected by a FITC-conjugated antibody to BrdU. The characteristic arc (Hoek and Stillman, 2003) of BrdU incorporation is present in the control cells, showing the high rate of incorporation in early S-phase, and only a small percentage of cells under the arc not incorporating BrdU. In contrast, the SLBP knockdown cells (S2) have a markedly lower BrdU-FITC signal overall and lack the characteristic arc observed in control cells (Figure 22B, bottom panel), suggesting that their rate of DNA synthesis is reduced. C2-treated cells treated with HU for 1 hr prior to addition of BrdU did not incorporate BrdU.

We also used the BrdU FACS analysis to quantify the cell cycle phase distribution at a higher resolution in the SLBP knockdown cells compared to the C2 control (Figure 22C). We quantified the G1 and G2 phases as the populations at 2N and 4N respectively that did not incorporate BrdU. We separated S phase into two phases, early and late, by quantifying the two populations that were between 2N and 3N, and 3N and 4N. It is clear from this representation (Figure 22C) that S2 cells accumulate both in the early part of S phase and also in late S-phase indicating the cells have difficulty completing S-phase. The number of cells in G2-phase was

reduced by 40%, consistent with cells progressing through S-phase more slowly but through G2/M and G1 normally.

Tom also determined the mitotic rate in the SLBP knockdown cells using an antibody to phosphorylated histone H3 (serine 10) as a mitotic marker. Positive cells were measured by FACS and were quantified as a percentage of the control (Figure 22D). It is clear from the mitotic index data that SLBP knockdown cells are going through mitosis at approximately half the rate of C2 control cells, consistent with the finding that G2 cells are reduced by 40%. This is in agreement with the interpretation that there is not a block in progression through G2, but rather a block in completing S-phase (and hence in entering G2).

Tom next determined whether the reduced rate of DNA replication resulted in activation of the S-phase checkpoint, a consequence of defective chromosome replication, damaged DNA, or environmental stress (Sancar et al., 2004). It has been shown previously that overexpression of a dominant negative chromatin assembly factor 1 (CAF1) or RNAi to CAF1 produces a similar DNA replication phenotype as knocking down SLBP (Hoek and Stillman, 2003; Ye et al., 2003). We hypothesized that reducing the levels of SLBP might result in impaired histone protein biosynthesis and hence mimic the defective chromatin assembly phenotype since the chromatin assembly machinery would not have enough histone proteins to properly assemble chromatin.

Tom assessed the phosphorylation of Chk1, a kinase that is activated by the ATR kinase when DNA replication stress is sensed in S-phase cells. As a control,



he again used C2 control cells treated with HU for 1 hour since HU treatment is known to activate the DNA replication checkpoint (Sancar et al., 2004). There was increased Chk1 phosphorylation on both serine 345 (S345) and serine 317 (S317) in response to SLBP knockdown (Figure 22E). The levels of Chk1 phosphorylation on each of these serine residues was higher than that in control cells but less than that in the HU-treated C2 siRNA population. This result supports the notion that reducing the levels of SLBP results in a partial inhibition of DNA replication phenotype that is directly sensed by the S-phase checkpoint sensor and effector molecules.

### **A minimal SLBP rescues the cell-cycle defect.**

SLBP functions in all aspects of histone mRNA metabolism from 3' end formation to degradation at the end of S-phase (Marzluff et al., 2008). I wanted to determine which of these functions was responsible for the accumulation of cells in S-phase when SLBP is knocked down. To this end, HA-tagged SLBP constructs were made which were mutant for translation (SAVEE) (Sanchez and Marzluff, 2002), RNA processing (9AA), and RNA binding (RR/KK) (Dominski et al., 2001), and resistant to the siRNA used to downregulate endogenous SLBP (Wagner et al., 2005). These constructs or a fully functional SLBP (Figure 23, top) which was resistant to the siRNA were cotransfected into cells along with the second dose of siRNA and cells were harvested for flow cytometry 72 hours later. Surprisingly, each of the mutants rescued the cell-cycle phenotype to the same extent as the fully functional SLBP (Figure 23, bottom) indicating that mutation of any single one of these functions is not sufficient to generate the S-phase accumulation. I next wanted to create a minimal SLBP capable of rescuing the cell-cycle phenotype and

generated multiple N- and C- terminal truncations to test in the assay. I was able to generate a piece of SLBP containing only the translation, RNA processing and RNA binding domains which rescued the phenotype (Figure 23, middle). Together these data indicate that the essential function(s) of SLBP lies in this region not in the first 64 or final 63 amino acids. I determined that this assay was not sensitive enough to diagnose the major lesion in histone mRNA metabolism that occurs when very low levels of SLBP are present and set out to characterize more fully the molecular phenotypes associated with the S-phase accumulation.

**Processed histone mRNA levels are only reduced 50% in SLBP knockdown cells.**

Since SLBP is involved in histone pre-mRNA processing, Tom determined the levels of histone mRNA, both by Northern blotting and by S1 nuclease mapping to assess whether the 3' end was formed properly. Tom used Northern blots to determine histone mRNA levels for all five histone proteins, and each band represents mRNAs from 5-15 genes (Marzluff et al., 2002) since the nucleotide sequences of the ORFs are very conserved. He used radiolabeled probes generated by PCR targeting the open reading frame (ORF) to detect histones H2A, H2B, H3, H4 and H1 and used 7SK snRNA as a loading control (Figure 24A). The abundance of histone H2A, H2B, H3, and H4 mRNAs is reduced to 37 to 70% of control levels in the SLBP knockdown cells when compared to the C2 control. This was a surprisingly high level, but it is consistent with previous reports (Zhao et al., 2004) which also reported only a relatively small reduction in core histone mRNA in HeLa cells.

Tom observed a slower migrating band in the histone H4 Northern blot (Figure 24A, panel 4), which was not present in the control cells. We hypothesized this represented misprocessed histone mRNA since if there were polyadenylation signals 3' of the genes the polymerase might read-through to that site. Tom therefore performed both S1 nuclease protection assays and Northern blots using total cell RNA and oligo(dT) selected total cell RNA to assess whether there were read-through mRNAs produced in the SLBP knockdown cells. We detected a small amount of misprocessed histone HIST2H2AA mRNA using S1 nuclease mapping, which mapped to a cryptic polyA site that is utilized when an intron is inserted into this gene (Pandey et al., 1990). There were no readthrough transcripts detected from the HIST2H3A gene. Tom used oligo(dT) selection to confirm that the slower migrating histone H4 mRNA was polyadenylated (Fig. 24C). Tom cloned the RNA from both the H2a and H4 histone genes where we observed misprocessing by oligo(dT)-primed RT-PCR to identify (1) the particular gene that was misprocessed and (2) the location of the polyadenylation site. Cloning of the histone H2A band using a pan-H2A primer targeting the ORF revealed this mRNA was derived from the HIST2H2AA gene and the polyadenylation occurred after the histone downstream element (HDE) (Figure 24D). The AATAAA in the H2A is conserved in the mouse and hamster (Debry and Marzluff, 1994) and the orthologue of this gene is polyadenylated using this canonical polyadenylation signal in rodent testis (Moss et al., 1994). It has been previously reported that this polyA<sup>+</sup> mRNA is present when NELF1 or SLBP is knocked down (Narita et al., 2007). Tom cloned the polyA<sup>+</sup> histone H4 mRNA and this mRNA is derived from the HIST1H4J gene. Two

different polyadenylation sites were found both of which use a polyadenylation signal present immediately 5' of the stemloop sequence (Figure 24E), which overlaps with the binding site for SLBP. This signal is not conserved in the mouse HIST1H4J gene. Interestingly, the first polyadenylated species resulted from cleavage and polyadenylation 5 nt 3' of the SL (21 nt from the AAUAAA), the same site where cleavage would occur in the presence of SLBP. The second species of polyadenylated HIST1H4J resulted from cleavage and polyadenylation 34 nt downstream of the AAUAAA signal, just prior to the HDE.

### **Histone protein levels are reduced in SLBP knockdown cells.**

Since there was still a substantial amount of properly processed histone mRNA in the SLBP knockdown cells for all the core histone proteins, we suspected that there must be a defect in histone protein production other than in histone pre-mRNA processing. Previous structure-function studies on SLBP have shown that there is a critical domain necessary for histone mRNA translation (Gorgoni et al., 2005; Sanchez and Marzluff, 2002). Due to the modest decrease in properly processed histone mRNA in SLBP knockdown cells, we hypothesized that reduction of SLBP might have had a more dramatic effect on histone protein synthesis than on histone pre-mRNA processing. I therefore investigated the levels of histone protein in SLBP knockdown cells. Figure 25D demonstrates that SLBP knockdown is >90%. I used Symplekin, a scaffold protein involved in 3' end formation (Takagaki and Manley, 2000b), as a loading control. Using an antibody against histone H3, I find that total histone H3 is reduced by about 50% in the SLBP knockdown cells, relative to the control cells.

To directly test the effect of knockdown of SLBP on the amount of histone protein, nuclear proteins were prepared from nuclei of control cells and cells with SLBP knocked down. The coomassie stained gel (Figure 25B) shows a reduction in total protein levels for all four core histones, consistent with the previous Western results for histone H3. Similar results have been found in five independent experiments, and the levels of the non-histone proteins remain constant. As judged by FACS analysis the DNA content per cell is the same in SLBP knockdown cells, which have an increased number of cells in S-phase. We conclude that the amount of histone protein produced in the SLBP knockdown cells is not enough to meet the histone protein demands of the replicating U2OS cell, resulting in a decrease in histone protein in the SLBP knockdown cells and this must, at least in part, contribute to the DNA replication and growth defects we observed.

#### **Regulation of histone mRNA degradation is impaired in SLBP knockdown cells.**

Histone mRNA half-life is coordinately regulated with the rate of DNA synthesis, and at the completion of S phase or upon inhibition of DNA synthesis there is a rapid degradation of histone mRNA. Histone mRNA degradation is a regulated process that requires translation of the histone mRNA (Graves et al., 1987; Kaygun and Marzluff, 2005b) and involves the addition of an oligo(U) tail to the 3' end in order to recruit specific decay factors for rapid, bidirectional decay of the mRNA (Mullen and Marzluff, 2008). Since reducing SLBP protein levels produced defects in DNA synthesis and normal histone gene expression, we asked if downregulation of SLBP had an effect on histone mRNA degradation in HeLa

cells. Tom inhibited DNA replication with 5 mM HU in both C2 control cells and SLBP knockdown cells for 30 and 60 minutes. A representative Northern blot analysis of histone H2A mRNA is presented in Figure 25C and reveals that the rate of histone mRNA degradation is decreased when SLBP protein levels are downregulated by siRNA. 7SK snRNA was used as a loading control to normalize the percentage of histone H2a mRNA remaining. Histone mRNA in the C2 control has been significantly reduced at the 30 minute time point and is barely detectable at the 60 minute time point (Figure 25C, lanes 2 and 3). In the SLBP knockdown cells, histone mRNA degradation is stabilized roughly 1.6 fold on average at the 30 minute time point and roughly 2.3 fold at the 60 minute time point (Figure 25C). The reduced rate of histone mRNA degradation could reflect the inefficient translation of histone mRNA with low levels of SLBP, and/or a requirement for SLBP in histone mRNA degradation.

**Processed histone mRNA is retained in the nucleus in SLBP knockdown cells.**

Another possibility for the low rate of histone protein synthesis is that the histone mRNA may be mislocalized. Eric and I determined the subcellular localization of the histone mRNA in SLBP knockdown cells. As a control experiment to localize histone mRNA, I synchronized HeLa cells using double thymidine block and released them into S phase for 3 hours and performed in situ hybridization with a probe to histone H2A. Histone H2A mRNA is localized strongly around the nucleus in a halo-like appearance, although some signal is present throughout the cytoplasm (Figure 26A, time 0). To assess probe specificity, I treated the synchronized cells with 5 mM HU for 15, 30, or 45 minutes since inhibition of DNA

synthesis results in rapid degradation of histone mRNA (Figure 26A, 15, 30 and 45 minute time points). As expected, histone H2A signal diminished over time and this provided confidence that our probe specifically detects the replication-dependent histone mRNAs.

Eric and I performed in situ hybridization to histone H2A mRNA in HeLa and H3 mRNA in U2OS cells in which SLBP has been knocked down. In the control HeLa cells, histone H2A mRNA was present in the cytoplasm, with the halo-like appearance in the preinuclear region with little to no staining in the nucleus (Figure 26B). In contrast, in the SLBP knockdown HeLa cells, in situ hybridization revealed that the histone H2A mRNA was concentrated in large foci in the nucleus, although some signal persisted in the cytoplasm. This phenotype was rescued by expressing an siRNA-resistant form of SLBP (SLBP\*, Wagner et al.) in HeLa cells (Figure 26B). We also performed in situ hybridization to histone H3 mRNA in U2OS cells where SLBP was knocked down (Figure 26C). In U2OS cells the phenotype is even more dramatic than in HeLa cells (Figure 26C). In U2OS cells, there is the characteristic halo-like staining of histone H3 mRNA in the C2 control treated cells but when SLBP is knocked down (S2), the histone H3 mRNA was concentrated in the nucleus (Figure 5C). A similar result was also observed in SLBP knockdown U2OS cells when in situ hybridization to histone H2a was performed (data not shown). Since the misprocessed RNA that is polyadenylated is presumably exported via the conventional polyA RNA export pathway, the histone mRNA which accumulates in the nucleus is properly processed mRNA.

As an alternative approach to confirm the presence of histone mRNA in the nucleus, Eric and I fractionated C2 control and S2 knockdown cells into nuclear and cytoplasmic fractions. We harvested protein and total RNA from these fractions as well as from a portion of the total cell lysate. Hsp90 and fibrillarin were used as cytoplasmic and nuclear markers respectively, and they were present almost exclusively in the appropriate fraction (Figure 24D). The first lanes in Figure 4D represent 100% of the total input from unfractionated cells. Similarly, the ribosomal RNA was largely in the cytoplasmic fraction (Figure 26E, lanes 2 and 3).

Eric and I also analyzed the histone mRNA in each of the subcellular fractions from the same experiment, and simultaneously probed for histone H3 mRNA and U1 snRNA as a control (Figure 26F). Consistent with the nuclear localization of spliceosomal snRNAs, we observed that U1 was concentrated in the nuclear fraction of both the C2 control and S2 knockdown (Figure 26F, lanes 2 and 5). A small amount of U1 snRNA was detected in the cytoplasm (Figure 26F, lanes 3 and 6). Histone H3 mRNA was concentrated in the cytoplasm of C2 control cells (Figure 26F, lane 3) but in the SLBP knockdown cells, about 50% of the histone mRNA was in the nucleus. Since the nuclear volume is about 10% of the volume of the cytoplasm (S2, Figure 26F, compare lanes 3 and 6), this result is consistent with the high concentration of histone mRNA in the nucleus detected by in situ hybridization. Sequestration of histone mRNA in the nucleus upon loss of SLBP is consistent with the decrease in histone protein synthesis due to decreased translation, as well as the failure to rapidly degrade histone mRNA when DNA replication was inhibited.



Taken together these results demonstrate that the major lesion in histone mRNA metabolism when SLBP is knocked down more than 95% is in export of the histone mRNA to the cytoplasm. Processed histone mRNA accumulates in the nucleus and is relatively stable. Thus the primary molecular defect in histone mRNA metabolism when SLBP is knocked down to 2-5% of normal levels is the retention of processed histone mRNA in the nucleus.

### **Tethering of SLBP to non-specific mRNA results in nuclear export.**

Many general mRNA export factors, including TAP and Rev, promote export to the cytoplasm when tethered to non-specific mRNAs (Yi et al RNA 2002 8 180). To determine if SLBP facilitates nuclear export of mRNA I utilized a well-characterized CAT (chloramphenicol acetyl transferase) reporter system. The pDM128/PL reporter contains a CMV driven bacterial cat gene located within an intron derived from HIV-1. Under standard conditions, the cat gene is spliced out of the reporter and not expressed due to its nuclear retention. Robust cat activity can be achieved by the insertion of 4 MS2 binding sites adjacent to the cat gene coupled with the expression of a heterologous fusion protein containing the MS2 coat protein and a factor capable of mRNA export (i.e. TAP). The activity results from the export of the cat mRNA prior to it getting spliced out of the reporter. 293T cells were cotransfected with either pDM128/PL or pDM128/PL-4XMS2 and HA-MS2 fused to TAP, Rev, or SLBP. Expression levels of MS2 fusions were assessed by Western blotting 48h after transfection, using an antibody to HA or SLBP. MS2-TAP and MS2-Rev resulted in increased levels of CAT activity from the pDM128/PL-4XMS2 reporter compared to the reporter lacking MS2 sites (Fig. 27A), as has been

reported for other transport factors (Wiegand et al., 2002; Yang et al., 2001).

Cotransfection of MS2-SLBP with the MS2 reporter results in CAT activity levels approximately 50% of those observed for TAP and Rev while coexpression of HA-MS2 had not effect on either reporter. Levels of MS2-SLBP, MS2-Rev, and HA-MS2 were detected by Western blotting using an antibody to MS2 (Fig. 27B). Note they levels of MS2-TAP expression were lower as reported previously (Wiegand et al., 2002). These data indicate that tethering SLBP to a non-specific mRNA is sufficient to promote nuclear export, albeit not as efficiently as TAP or Rev.

## **DISCUSSION**

As the only mRNAs that are not polyadenylated, metazoan histone mRNAs require a distinct set of factors for pre-mRNA processing, translation and regulation of stability. A common factor in all these processes is SLBP, which binds the stemloop in the histone pre-mRNA during transcription, remains associated with the mRNA during translation and is involved in histone mRNA degradation (Marzluff et al., 2008). Here we have ascertained the effect of knocking down SLBP in mammalian cells to less than 5% of the normal level. The cells remain viable and continue growing at a slow rate. Surprisingly there is only a modest effect on the accumulation of processed histone mRNA, but a substantial fraction of the histone mRNA is retained in the nucleus, suggesting a previously unrecognized role of SLBP in histone mRNA export in vivo.

We and others have previously shown that knocking down SLBP results in slow growth of cells, and an accumulation of cells in S-phase(Wagner et al., 2005; Zhao

et al., 2004). This effect is due to the SLBP knockdown since normal growth can be restored by expression of an RNAi resistant SLBP (Wagner et al., 2005)..

Progression through S-phase was slowed as judged by the release of cells from either nocodazole or HU blocks (Zhao et al., 2004). This is in contrast to knocking down components of the U7 snRNP, which result in an arrest of cells in mid-G1 (Wagner and Marzluff, 2006), suggesting that the cell monitors the ability to synthesize histone mRNA by monitoring the presence of an active U7 snRNP, but that cells enter S-phase normally when SLBP is reduced.

SLBP is a relatively unstable protein, being degraded at the end of S-phase and not accumulating again until just prior to entry into the next S-phase (Whitfield et al., 2000; Zheng et al., 2003). Thus it is relatively easy to obtain effective knockdown of the SLBP protein, although for the phenotypes observed here SLBP must be knocked down >95%, and with long exposures of the Western blots we always detect small amounts of SLBP in the knockdown cells. Thus the results reported here reflect the phenotypes observed when SLBP is severely limiting but not absent from cells.

The most obvious effect of knocking down SLBP is a slow growth rate, and a change in cell cycle distribution of the cells, with an accumulation of cells in S-phase. The cells continue dividing and we see no evidence of apoptosis during the 5-6 day course of the experiments. The accumulation of cells in S-phase is due to a slower progression through S-phase as a result of a reduced rate of DNA replication (Fig. 22). We suggest that the reduced rate of replication is due to a failure to provide replication-dependent canonical histones to assemble chromatin at a high rate. This

likely results in defects at the replication fork, leading to activation of the S-phase checkpoint, slowing the rate of DNA replication. Similar phenotypes have been observed with knockdown of CAF1, the major chromatin assembly factor for the canonical, but not the replication-independent histones (Ye et al., 2003).

Surprisingly the overall level of histone mRNAs is only reduced about 2-fold in SLBP knockdown cells, and the great majority of the histone mRNA is accurately processed. This implies that the residual SLBP in the cell must be preferentially used for histone pre-mRNA processing. One possibility is that the residual SLBP might be concentrated in the Cajal body, where it could function in histone pre-mRNA processing, but we were not able to detect a specific localization by immunofluorescence. There is a small amount of misprocessed polyadenylated mRNA that accumulates from some of the histone genes, suggesting that the efficiency of processing is reduced in the SLBP knockdown mutants. The HIST2H2AA gene has been previously shown to have a cryptic polyadenylation site in the HDE which is used when histone processing is disrupted by insertion of an intron (Pandey et al., 1994), and in round spermatids (Moss et al., 1994). This mRNA is also polyadenylated when NELF, a transcription elongation inhibitor, is knocked down (Narita et al., 2007). There is also a polyadenylated histone H4 mRNA that accumulates when SLBP is knocked down. This mRNA is derived from a single histone H4 gene, HIST1H4J, which contains a polyadenylation signal in the 3' UTR, just before the stemloop. Although the position of the polyA tail is at the same site as the cleavage site for histone mRNA, it is likely that this polyadenylation signal, which overlaps the SLBP binding site, directs the cleavage of the HIST1H4J

pre-mRNA to yield the polyadenylated mRNA. Each of these polyadenylation signals overlaps a cis-element required for histone mRNA processing, suggesting that normally the histone processing machinery preferentially associates with these signals.

Unlike the situation in *Drosophila* (Godfrey et al., 2006; Lanzotti et al., 2002), there is not extensive readthrough of histone transcripts resulting in polyadenylated mRNAs in mammalian cells when SLBP is reduced. This may reflect a need in *Drosophila* for efficient termination between closely spaced histone genes, or there may be a role for polyadenylated histone mRNAs in some cells at some stage. The role of SLBP in the processing of histone pre-mRNA may be “catalytic” allowing cells to perform this function, even under limiting concentration of SLBP, and the SLBP can then be released for processing more histone mRNAs. In contrast the requirement for SLBP in other steps in histone mRNA metabolism (e.g. export or translation) is likely stoichiometric and the RNAi-mediated reduction in SLBP levels would interfere with these functions.

**Transport of histone mRNA.** Most mRNAs are transported through a pathway that requires TAP (yeast mex67p) as an adaptor protein for export (Erkmann and Kutay, 2004). TAP can be recruited to mRNAs derived from spliced pre-mRNAs by binding to components of the exon-junction complex deposited near the exon-exon border after splicing. Unspliced mRNAs, such as histone mRNAs, must recruit TAP by a different mechanism. There have been several studies relevant to histone mRNA export from the nucleus. In one study the role of processing was assessed by forming the histone 3' end either by a normal processing reaction or by ribozyme

cleavage (Eckner et al., 1991). They found that the complete histone mRNAs were cytoplasmic regardless of the mechanism of 3' end formation. However, if the histone ORF was replaced with the chloramphenicol acetyl transferase ORF, then the mRNA formed by the processing reaction was cytoplasmic, while the mRNA formed by the ribozyme was retained in the nucleus.

Huang and Steitz studied elements in the histone ORF that were capable of directing transport of an heterologous mRNA in *Xenopus* oocytes and identified a binding site for the SR protein 9G8 responsible for this activity (Huang and Steitz, 2001), which is in a region of the histone ORF capable of conferring cytoplasmic localization on other unspliced mRNAs (Huang and Carmichael, 1997). These studies suggest a role for the 9G8 protein in recruiting TAP, and provides a partial explanation for the observed conservation in the nucleotide sequence level in histone ORFs (Debry and Marzluff, 1994; Friend et al., 2007). In a previous study of histone mRNA export in *Xenopus* oocytes, we found that there was not a requirement for any specific nucleotide sequence, including the 3' end of histone mRNA or SLBP, for histone mRNA export (Erkman et al., 2005a) in *Xenopus* oocytes. This study suggested that there are multiple potential binding sites for proteins that could recruit TAP, and that in *Xenopus* oocytes there is a large supply of these proteins, in agreement with other studies which showed that the length of the RNA was a major determinant of its transport efficiency in *Xenopus* oocytes (Masuyama et al., 2004).

In mammalian cells with active RNA metabolism, there is likely to be competition among the mRNAs for the binding proteins and export factors, just as there is for factors involved in RNA processing. Thus it is not surprising to find that there might

be more stringent requirements for export in mammalian cells than in *Xenopus* oocytes. Our results suggest that there is a critical role for SLBP in histone pre-mRNA export in mammalian cells.

Overall these results are consistent with a model where multiple factors may need to bind to an RNA for efficient export. The processing reaction may facilitate recruitment of some of these factors (Eckner et al., 1991) and there are clearly sequence elements in the histone ORF that promote export (Huang and Carmichael, 1997; Huang and Steitz, 2001) of mRNAs. Our results suggests that association of SLBP is necessary to promote efficient transport of histone mRNA, but not to stabilize the histone mRNA in the nucleus.

When SLBP levels are greatly reduced, mature histone mRNAs accumulate in the nucleus. A striking aspect of this result is that these mRNAs accumulate to substantial levels. When mutations are introduced into the histone stem-loop that alter SLBP binding (Pandey et al., 1994) or the HDE is deleted preventing r U7 snRNP binding (Chodchoy et al., 1991), or when the HDE is naturally suboptimal (Liu et al., 1989), then the unprocessed mRNAs do not accumulate, but rather must be rapidly degraded. The same is true of improperly processed mRNAs from polyadenylated transcripts. In the SLBP knockdown cells histone mRNAs are properly processed, and direct sequencing of the mRNAs after circular RT-PCR (Mullen and Marzluff, 2008) did not reveal any alterations at the 3' end of the mRNA (EJW, unpublished). Thus these mature mRNAs, which are not bound to SLBP, are still stable in the nucleus. They may be stable in the nucleus since they have gone through the normal processing pathway, even though they can not get exported. A

model for histone mRNA export in normal cells and for the nuclear retention of histone mRNA in SLBP knockdown cells is shown in Fig. 26.

### **Decreased levels of histone protein in SLBP knockdown cells.**

The second feature of the SLBP knockdown cells is that they have a lower amount of histone protein per cell, consistent with a defect in histone protein biosynthesis. Initially we found this surprising since one might think that reducing histone protein:DNA ratios 2 fold might have a profound effect on cell physiology and viability. However, in at least one situation, infection of human T lymphocytes with HTLV1, there is an approximately 2-fold reduction in histone:DNA ratios, which is mediated by a reduction in core histone mRNA expression without affecting DNA replication (Bogenberger and Laybourn, 2008). In yeast one can vary the histone:DNA ratio by changing the dosage of the histone genes (Gunjan and Verreault, 2003) or in some mutants in chromatin modification (Xiao et al., 2007) and personal communication from B.D. Strahl) without affecting viability.

### **Conclusions.**

SLBP is essential for the proper accumulation of histones for assembly of newly replicated DNA into chromatin. Our results suggest that a major consequence of SLBP depletion is the failure to export properly processed histone mRNA. This suggests that SLBP must be associated with histone mRNA for export, and that SLBP may dissociate from histone mRNA after processing and additional nucleoplasmic SLBP may then reassociate with the histone mRNA prior to export. Understanding and identifying the components of the nuclear histone mRNP, as well as the cytoplasmic histone mRNP, will help elucidate the pathway between

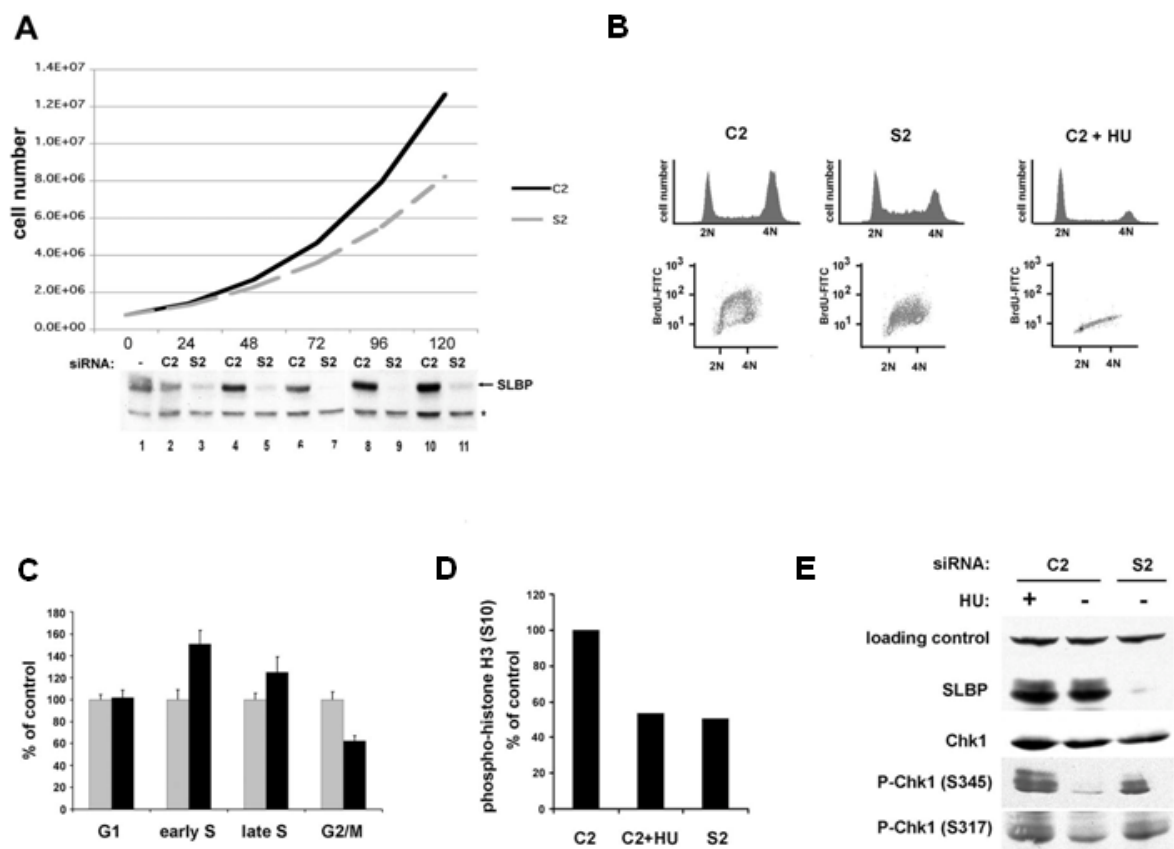


processing of the mRNA at the site of transcription, and subsequent export and translation of the histone mRNA.

**Figure 22. SLBP knockdown cells are viable but grow at a slower rate as a result of various S phase defects.** (A) Growth curve (total cell number) of control (C2) and SLBP (S2) RNAi populations. Lysates from the timepoints shown in the graph were probed for levels of SLBP (\*-represents cross-reacting band).

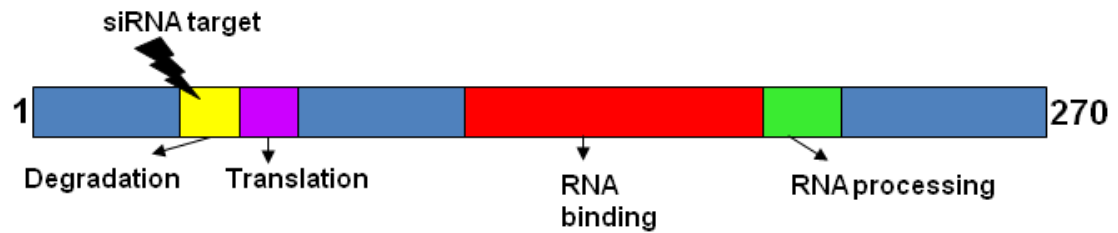
(B) SLBP knockdown results in S-phase accumulation and a reduced rate of DNA replication. C2 and S2 treated U2OS cells were pulsed for 1 hr with 10  $\mu$ M bromodeoxyuridine (BrdU) and harvested and analyzed by FACS for DNA content and DNA synthesis. Hydroxyurea (HU, 5 mM) was added to a parallel culture of C2 cells 1 hr prior to the addition of BrdU. Cells were fixed and BrdU incorporation detected with an anti-BrdU antibody conjugated to FITC and for DNA content with propidium iodide. Cells were analyzed using a FACScan device and Summit software for DNA content and intensity of BrdU incorporation. (C) The cell cycle distribution (G1, early and late S phase, and G2/M) was quantified from the data in (B), (C2, grey bars; S2, black bars) from 3 independent experiments. Quantities are expressed as a percentage of the C2 control cells. (D) Parallel cultures from (B) were fixed, stained with an antibody to the mitotic marker phospho-histone H3, and the antibody detected with a secondary antibody conjugated to FITC. Cells were analyzed similarly as in (B) and the population of cells with fluorescent signal in the G2 population was quantified. The number of phosphor-histone positive cells is expressed as a percentage of the C2 control cells. (E) (E) SLBP knockdown results in the activation of a Chk1-dependent cell cycle checkpoint. Protein lysates from SLBP knockdown cells were analyzed for the predominant signal transduction pathway that responds to replication stress by Western blotting for SLBP, total Chk1

protein and phospho-specific Chk1 (serines 317 and 345) was analyzed by Western blot analysis. HU was used a positive control for Chk1 activation.

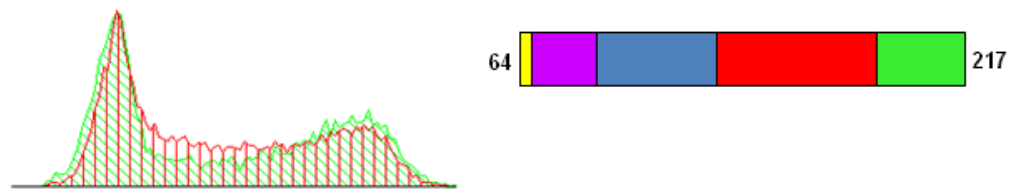


**Figure 23. A minimal SLBP construct rescues the cell-cycle phenotype associated with SLBP knockdown.** (Top) Schematic of full-length SLBP showing the locations of degradation, translation, RNA binding and RNA processing domains as well as the siRNA target site. (Middle, right) Schematic of Mini2-SLBP showing domains included in this truncation. (Middle, left) FACS profiles of SLBP knockdown cells (red) and Mini2-SLBP rescued cells (green). (Bottom) Table showing cell cycle distributions of rescues by various mutants.

## Full-length SLBP



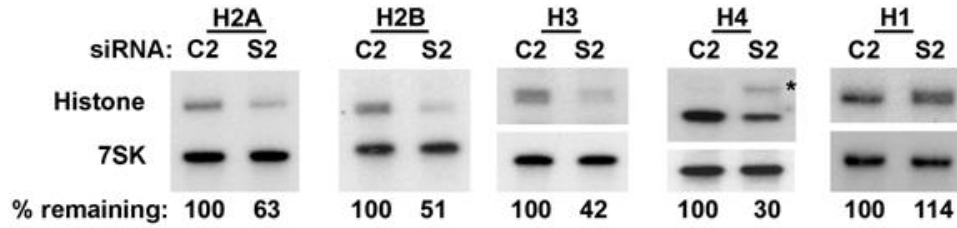
## Mini2



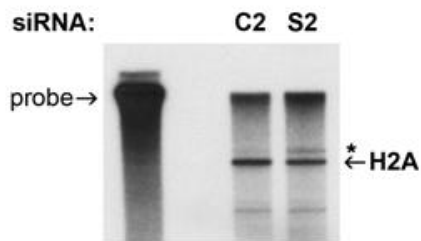
siRNA	Plasmid	G <sub>1</sub>	S	G <sub>2</sub>	Mutated Function
C2	GFP	37.4	31.8	31.7	
S2	GFP	38.5	40.8	21.4	
S2	hSLBP	46.3	29.2	25.1	
S2	Mini- SLBP2	42.9	31.3	25.8	
S2	SAVEE	46.8	30.0	23.0	Translation
S2	9AA	44.9	31.0	24.1	Processing
S2	RR/KK	49.0	29.0	22.0	RNA binding

**Figure 24. Histone mRNA levels are reduced in SLBP knockdown cells and a portion of the mRNA is misprocessed and polyadenylated.** (A) Northern blot analysis was performed on 2  $\mu$ g of total RNA from C2 and S2 treated cells. 7SK snRNA is used as a loading control. The percentage of histone mRNA present in the S2 treated cells was quantified using PhosphorImager analysis and ImageQuant software and normalizing mRNA to the 7SK snRNA levels. These results are representative of several different experiments. Note that a slowly migrating product reacting with the histone H4 probe (\*) is present in the S2 cells. (B) 3' S1 nuclease protection assays were used to map the 3' ends of histone H2A (HIST2H2AA) (lanes 2 and 3) and histone H3 (HIST2H3A/C) (lanes 5 and 6) mRNAs. The asterisk (\*) protected fragment in lane 3 indicates a longer mRNA which is polyadenylated (panel D).. (C) Total RNA was selected on oligo(dT) cellulose and 2 micrograms of total RNA (lanes 1 and 2) and polyA + RNA (lanes 3 and 4) were analyzed by Northern blotting for histone H4 mRNA. The asterisk (\*) indicates the longer polyadenylated H4 mRNA. (D and E) Total RNA from C2 and S2 U2OS cells were primed for reverse transcription with oligo(dT) fused to a T7 adaptor sequence and subjected to PCR using a pan-H2A ORF oligo (panel D) or a pan-H4 oligo (panel E) together with a T7 adaptor sequence. The products were analyzed by gel electrophoresis, and the products were cloned and multiple clones from each reaction sequenced.

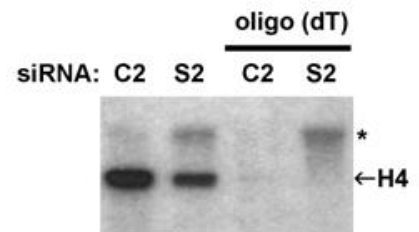
**A**



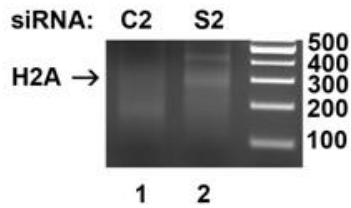
**B**



**C**



**D**



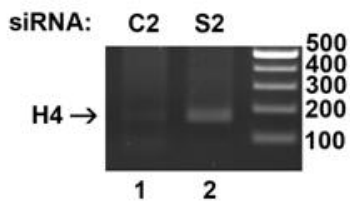
*HIST2H2AA* 3' UTR:

CGACCGGCCCGCGTCTCGAAGGGCACCTGTGAAC  
TCAAAAGGCTCTTTTCAGAGCCACCCACGTTTCA  
AATAAAGAGTTGTTAATGCTAAAAAAAAAAAAAAAAA  
poly(A) HDE

*HIST1H4J* 3' UTR:

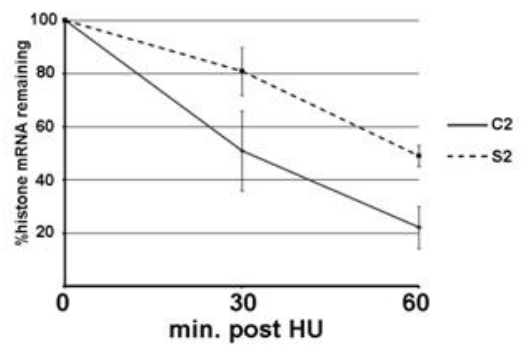
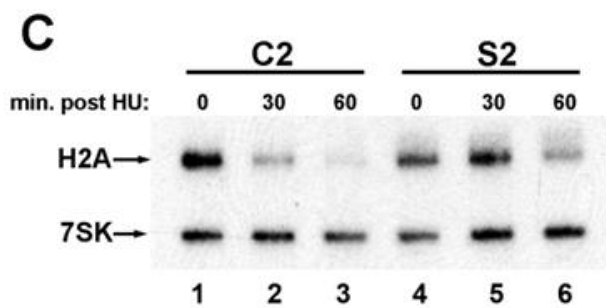
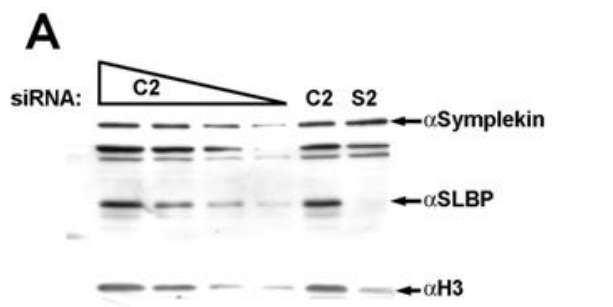
GGGCCGCACCCTCTACGGTTTCGGTGGTTGAGCGTC  
CTTTTCTACCAATAAAAGGCCCTTTTCAGGGCCACCC  
TAAAAAAAAAAAAAAAAAAAAAAAAA  
GGGCCGCACCCTCTACGGTTTCGGTGGTTGAGCGTC  
CTTTTCTACCAATAAAAGGCCCTTTTCAGGGCCACCC  
TACTTTCTCAGCTGAAAAAAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAA

**E**



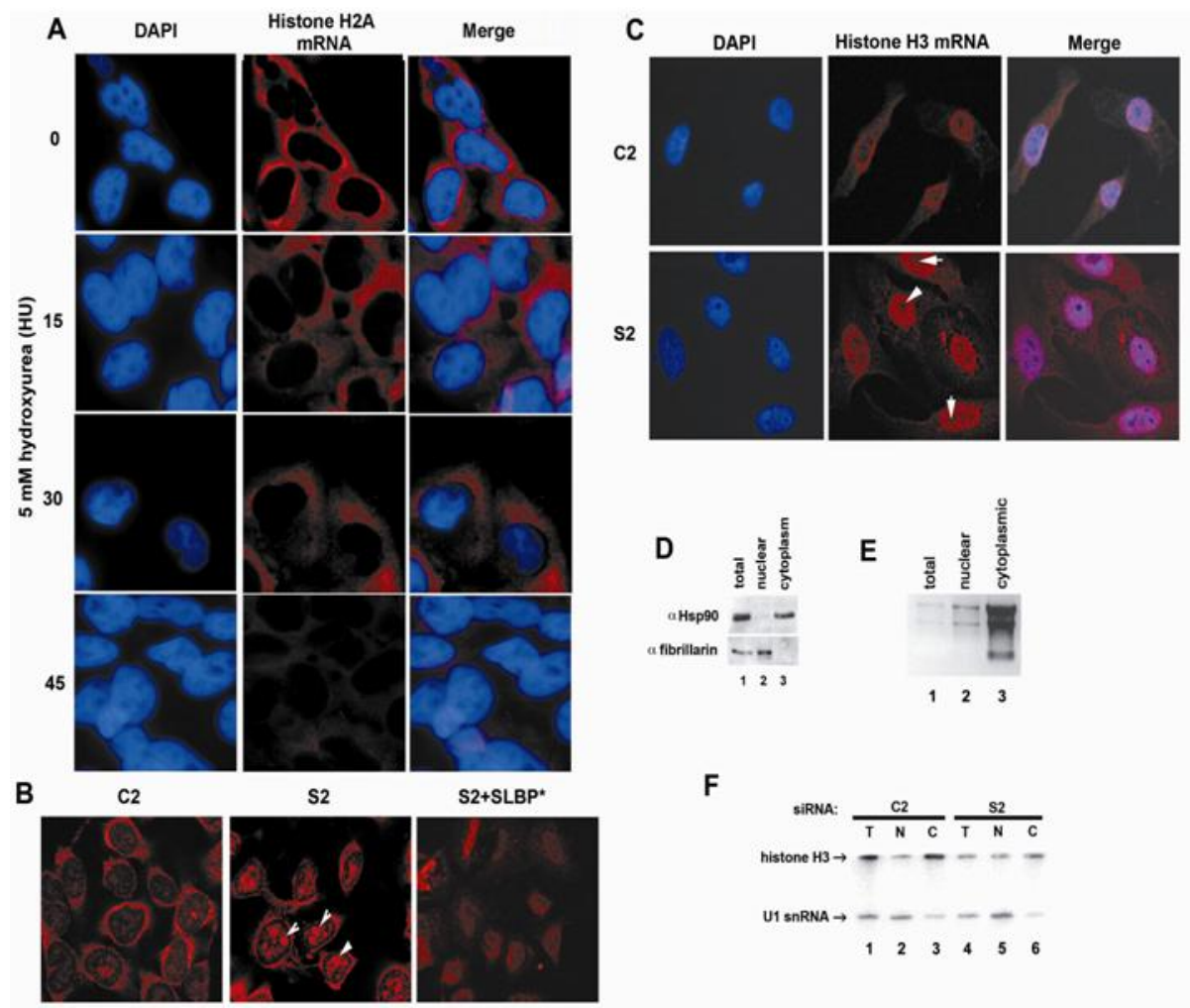


**Figure 25. Histone protein levels are decreased in SLBP knockdown cells and the normal histone mRNA degradation in response to HU is impaired.** (A) A serial dilution Western blot analysis of C2 (lanes 1-4) and a direct comparison of C2 (lane 5) and S2 (lane 6) total cell protein lysates. The levels of SLBP, histone H3 protein and Symplekin as a loading control were analyzed. (B) U2OS cells were knocked down for SLBP and nuclear proteins prepared from the control and SLBP knock down cells. They were resolved on 15% SDS-polyacrylamide gels and stained with Coomassie Blue. The core histone proteins are indicated (arrows) (C) Histone H2A mRNA is not degraded as efficiently in SLBP knockdown cells compared to control cells. HeLa cells were treated with C2 (lanes 1-3) or S2 (lanes 4-6) siRNAs, treated with 5 mM HU and cells harvested 30 (lanes 2 and 5) or 60 min (lanes 3 and 6) later. Total cell RNA was analyzed by Northern blotting (left panel) with probes to H2A mRNA and 7SK snRNA as a loading control. The right panel represents the results from 3 independent experiments. Standard deviations are indicated by vertical bars.

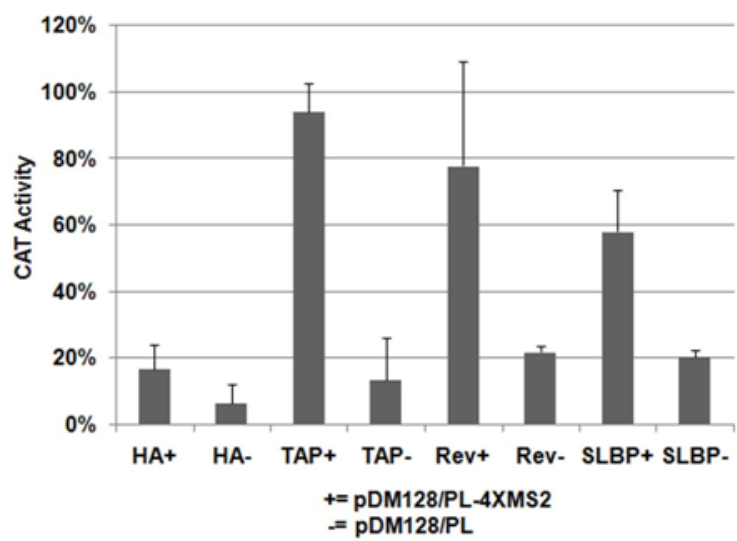
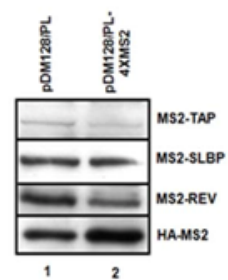


**Figure 26. Histone mRNA in SLBP knockdown cells is retained in the nucleus.**

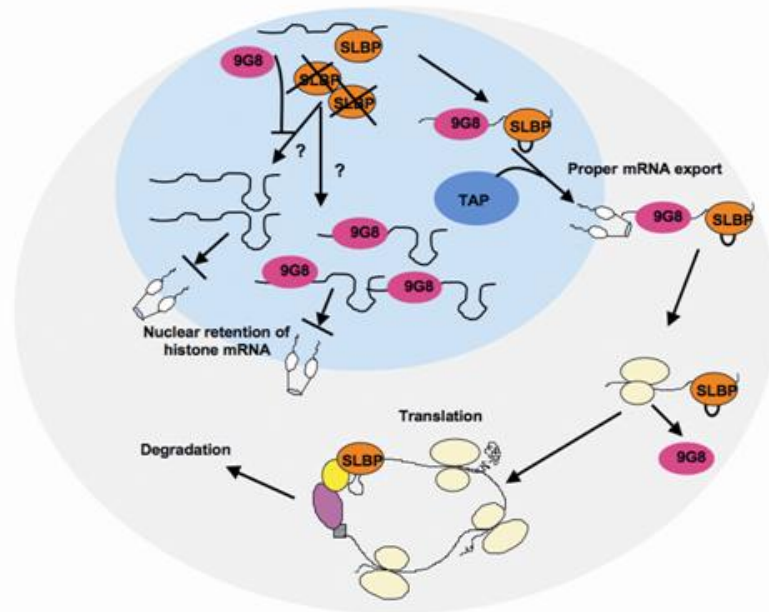
(A) HeLa cells synchronized by double-thymidine block were released into S phase for 3 hours and treated with 5 mM hydroxyurea (HU) for 0, 15, 30 and 45 minutes. Cells were fixed and prepared for in situ hybridization with antisense DIG-labeled H2a mRNA. Cells were stained with DAPI and anti-DIG antibodies. DAPI (blue), histone H2A mRNA (red). (B and C) Histone H3 mRNA is localized to the nucleus when SLBP is knocked down. (B) Logarithmically growing HeLa cells were treated with the indicated siRNAs and in situ hybridization performed with antisense DIG-labeled H3 mRNA. Note the histone mRNA accumulates in nucleoli of HeLa cells in S2 treated cells (arrows). (C) U2OS cells were treated with the indicated siRNAs and in situ hybridization performed with antisense DIG-labeled H3 mRNA as in (B). Cells were stained with DAPI. Note that there is no accumulation of histone mRNA in nucleoli (arrows) in U2OS cells. (D and E) U2Os cells were fractionated into nuclear and cytoplasm and protein (panel D) or RNA (panel E) from total and subcellular fractions analyzed. Total (lane 1), nuclear (lane 2), and cytoplasmic (lane 3) protein fractions were separated by 12% polyacrylamide gel electrophoresis and probed for Hsp90 (top, as a cytoplasmic marker) and fibrillarin (bottom, as a nuclear marker). Proteins were analyzed by Western blotting for Hsp90, and fibrillarin markers. (E) Total RNA was analyzed by Northern blot analysis and histone H3 mRNA and U1 snRNA (nuclear marker) simultaneously detected using a mixture of radiolabelled probes.



**Figure 27. SLBP promotes mRNA export. A)** 293T cells were transfected with 25ng of reporter plasmid and 500 ng of MS2 fusion protein plasmid. 48h post transfection, cells were harvested in TEN buffer and analyzed by western blotting. **B)** CAT assay was performed on lysates from A. CAT activity was normalized to untransfected lysates and the maximum value for each data set was set to 100% CAT activity. Values are for 3 independent experiments with error bars of 1 standard deviation.

**A****B**

**Figure 28. Model of SLBP-dependent histone mRNA export.** **A.** Histone mRNA metabolism in normal S-phase cells. SLBP remains associated with the histone mRNA in the nucleus and the mRNA is efficiently exported associated with multiple proteins that recruit TAP. **B.** Histone mRNA metabolism in cells depleted of SLBP. Histone mRNA is processed normally by the residual SLBP but that SLBP dissociates from the processed mRNA before it can be exported, and the processed message remains in the nucleus where it is stable.





## **Chapter 5**

### **Summary and Conclusions**

#### **Introduction**

Evidence abounds from yeast (Meeks-Wagner and Hartwell, 1986) to *Drosophila* (Sullivan et al., 2001) to mammals (Arnold et al., 2008) that proper regulation of histone gene expression is essential. Metazoan histone mRNAs have evolved a unique mechanism of regulation as a result of the novel 3' end on these mRNAs. These histone pre-mRNAs contain no introns and their 3' ends consist of a stemloop structure rather than a poly(A) tail. Thus maturation of histone mRNAs requires only a single endonucleolytic cleavage 3' to the stemloop. This 3' end is bound by SLBP and the SLBP-stemloop complex remains active throughout the life of the histone mRNA and performs the essential functions of the poly(A)-PABP complex on other mRNAs. I will summarize the findings of my dissertation and place them into context current knowledge of mRNA 3' end formation and histone mRNA metabolism in general.

In chapter 2 of this dissertation, I showed that proper 3' end formation of histone mRNA in *Drosophila* requires a subset of polyadenylation factors and that this subset exists as a “holo-cleavage factor” which is active in both histone and poly(A) containing mRNA 3' end formation. Chapter 2 also shows that processing of histone pre-mRNAs occurs cotranscriptionally, and that factors required for 3' end formation are recruited to histone genes in vivo. Chapter 3 examines the mechanism of transcription termination of histone genes in mammals and reveals that, due to the U7 snRNP dependence of 3' end formation and the surprising finding that CPSF73 has both endo- and exonucleolytic activities, termination of transcription is not Xrn2 dependent as it is for poly(A) containing genes. And, finally, in Chapter 4, I describe a new role for SLBP in histone mRNA metabolism. This dissertation reveals some interesting new findings in the area of the regulation of histone gene expression and opens up some exciting new lines of questioning for future studies.

**A subset of cleavage and polyadenylation factors form a “holo-cleavage factor” active in 3' end formation of all mRNAs.**

Transcription of protein coding genes yields large RNA species termed pre-mRNAs which must be processed in multiple steps to yield mature mRNAs, which are capable of being exported and translated. Processing steps include addition of a modified cap to the 5' end, splicing out of introns, and proper 3' end formation. For all mRNAs except histone mRNAs, 3' end formation consists of endonucleolytic

cleavage and addition of a long (~250 nt) poly(A) tail. Cleavage and polyadenylation is a complex process which is guided by multiple cis-acting elements including the poly(A) sequence (AAUAAA) and a G/U rich DSE and carried out by a large protein complex containing as many as 85 polypeptides (Shi et al., 2009). The PAS is located 5' to the cleavage site (usually a CA dinucleotide) and is bound by the CPSF complex, while the DSE, located 3' to the cleavage site, is bound by the CstF complex. These complexes act in concert to recruit the cleavage factor and polyadenylation machinery and position them at the correct site. Histone mRNAs have neither a PAS nor a DSE; they do, however, contain analogous cis-elements in the stemloop and histone downstream element. These elements are bound by SLBP and the U7 snRNP respectively, which act to recruit and position the cleavage machinery properly.

Surprisingly, several components of the poly(A) machinery are shared with histone mRNA processing. CPSF73 was shown to be the cleavage factor for both histone (Dominski et al., 2005b) and poly(A) (Mandel et al., 2006) containing mRNAs and Symplekin was also shown to be involved in both processes (Kolev and Steitz, 2005; Takagaki and Manley, 2000a). CPSF100, long known to be a critical component of the polyadenylation apparatus is also required for histone pre-mRNA processing in *Drosophila* (Wagner et al., 2007) and mammals (Kolev et al., 2008). The finding that there are common components in the two pathways led me to ask how these factors interact with one another and with the accessory proteins required for their respective processes

### **Stable complexes containing polyadenylation factors in *Drosophila*.**

Codepletion experiments in *Drosophila* revealed that CPSF73, CPSF100 and Symplekin exist as a stable complex because depletion of any one of these factors resulted in strong downregulation of the other two, and the three proteins were efficiently coimmunoprecipitated. Knockdown of any of these factors affects both histone and poly(A) mRNA processing; resulting in poly(A) site switching on a poly(A) containing mRNA and polyadenylation of histone mRNAs. Notably, histone misprocessing after knockdown of these components of the holo-cleavage factor is different than in knockdown of a histone specific factor such as SLBP. While polyadenylation does occur, the majority of the misprocessed RNAs are very long and not polyadenylated indicating that removal of these proteins affects both histone and poly(A) cleavage site usage generating long readthrough transcripts. These experiments also gave us our first hint that there are discrete complexes containing the holo-cleavage factor. Knockdown of CPSF160 results in codepletion of Symplekin, CPSF73 and CPSF100, yet has no effect on histone 3' end formation. Similarly, CstF64 downregulation results in Symplekin depletion but does not affect histone pre-mRNA processing.

I also wanted to determine how CPSF73, CPSF100 and Symplekin interact with one another as well as how they interact with accessory factors, in this case, the histone-specific processing factors SLBP and Lsm11. IP experiments under stringent detergent conditions confirmed that CPSF73, CPSF100 and Symplekin are indeed tightly bound to one another as they reciprocally coIP one another with high efficiency. Attempts to coIP these components of the cleavage factor with SLBP or Lsm11 were unsuccessful suggesting that the histone specific machinery is not a

stable component of the cleavage factor. I was able to coimmunoprecipitate CPSF73 and Symplekin with SLBP and Lsm11 under low detergent conditions indicating that this interaction is weaker and likely more transient. These data supports the notion that there is a core cleavage factor which may associate with proteins either from the histone or poly(A) specific processing apparatus, making discrete cleavage factors. Further supporting this idea is the fact that CstF77 does not coIP with SLBP or Lsm11 or with CPSF73 or Symplekin. Knockdown of this factor has no effect on histone pre-mRNA processing and the fact that it does not interact with histone-specific factors demonstrates that independent complexes may exist. Many of the components of the polyadenylation machinery have been identified, but clearly some factors required for histone pre-mRNA processing remain to be identified.

The revelation that much of the cleavage machinery is conserved between histone and poly(A) mRNAs raises some interesting points. In mammals, the cell cycle regulation of histone biosynthesis occurs largely at the level of RNA processing as histone mRNA levels are high in S-phase due to an increase in processing efficiency. SLBP is a primary cell cycle-regulated factor required for proper histone mRNA 3' end formation as the U7 snRNP is stable throughout the cell cycle. When SLBP is synthesized, the cell is already poised to produce large amounts of histone mRNA since the other required components are already present. In *Drosophila*, the mechanism using different cis-elements to generate mature 3' ends is conserved. Therefore, higher eukaryotes have evolved a method of cleaving histone mRNAs that conserves some of the polyadenylation machinery, eliminating

the need for an entirely separate mechanism. These complexes must contain unique factors to direct the core cleavage complex to the proper pre-mRNA. CPSF160 likely plays this role in poly(A) mRNAs and the factor that carries out this function for histone pre-mRNAs is not known.

### **Histone 3' end formation occurs cotranscriptionally.**

Fidelity in RNA processing is essential for proper gene expression and regulation. To this end eukaryotic cells intergrade many RNA processing events with transcription, such that as nascent RNA is extruded from RNA Pol II, it is acted upon by RNA processing factors responsible for capping, splicing, and 3' end formation. Indeed, 3' end formation serves to release the transcript from the chromatin template. In the case of poly(A) mRNAs, this cotranscriptionality is achieved by recruitment of factors to the RNA by the phosphorylated CTD of RNA Pol II (Proudfoot, 2004). However, the phosphorylation pattern of the CTD is distinct in Cajal bodies as it lacks Ser2 phosphorylation (Xie and Pombo, 2006), the modification which recruits the polyadenylation machinery (Licatalosi et al., 2002b). Initially, I wanted to determine whether or not processing of histone mRNAs was cotranscriptional. When I began my dissertation work there was no direct evidence for cotranscriptionality and, in fact, the only paper concerning the cotranscriptional processing of histone pre-mRNAs posited that cotranscriptional processing was not more efficient than processing of a synthetic pre-mRNA in vitro (Adamson and Price, 2003), in contrast to findings with poly(A) mRNAs (Adamson et al., 2005). During the course of my studies, two groups published papers demonstrating that indeed

some factors such as CPSF73 are recruited to mammalian histone genes (Glover-Cutter et al., 2008; Narita et al., 2007).

I demonstrate in chapter 2 that Symplekin and CPSF73 associate with histone genes in vivo using ChIP assays. These proteins are localized across the histone gene locus. The small size of the histone genes (~400 nt) approaches the limit of ChIP resolution using sonication which generates 200-500 bp fragments of DNA making it impossible to accurately determine when they associate with the gene. These factors are not located in the 1.3 kb intergenic region between the H3 and H1 genes, indicating that they are indeed recruited to the genes themselves and not just in close proximity to the cluster. The absence of RNA Pol II in this intergenic region also indicates that termination of transcription occurs close to the cleavage site, an observation I will discuss in the next section. Notably, CstF50, a polyadenylation factor not involved in histone pre-mRNA processing associates with histone genes only at low levels relative to those with which it associates with a poly(A) containing gene. CstF50 does appear to associate with histone genes at levels slightly above background, which may indicate that even with the proposed modular nature of the cleavage complex, sometimes factors not required for processing of a particular RNA may end up associated with those genes. CstF50 is, however, strongly recruited to histone genes when polyadenylation is activated by depletion of SLBP. These data demonstrate that the cleavage complex is capable of switching from its histone mRNA processing form to a poly(A) processing form rapidly when necessary. It may be that the CstF50 associated with histone genes under normal conditions serves as a failsafe sensor, that is, upon activation of

polyadenylation, some of the required machinery is present which might serve to recruit additional help or it may represent background association with RNA Pol II.

Further ChIP data presented in Chapter 2 support the hypothesis that a core cleavage factor acts in both histone and poly(A) pre-mRNA processing independently from one another. Knockdown of CPSF160 has no effect on histone pre-mRNA processing, but does deplete Symplekin, CPSF73, and CPSF100. When I ChIP with antibodies to Symplekin or CPSF73 in CPSF160 knockdown cells, I found a specific depletion of these factors from canonical poly(A) containing genes, but not from histone genes. I hypothesize that this is due to depletion of only that subset of the holo-cleavage factor which is associated with CPSF160, and not that active in histone processing. Knockdown of CstF64 has a similar affect on recruitment of Symplekin to poly(A) genes without affecting histone genes. Thus these data support the existence of two distinct cleavage factors.

ChIP studies using antibodies to histone-specific processing factors will need to be performed to enhance our understanding of the relationship between transcription and 3' end formation. One of the major questions that remains regarding the cotranscriptional nature of histone pre-mRNA processing is how these factors are recruited to the genes. Several possibilities exist including; high local concentration of processing factors in HLBs, direct recruitment by RNA Pol II, or a combination of these. Histone-specific processing factors could also be recruited by transcription factors which localize to histone genes such as NPAT. The argument that high local concentration is important is an attractive one as many of the factors involved in histone processing are concentrated in HLBs including Symplekin and



U7 snRNP, but this does not account for the fact that these factors are absent from an intergenic region by ChIP analysis. Recruitment via interaction with the CTD of RNA Pol II is a common mechanism for recruitment of RNA processing factors to sites of activity, but the common modification used to recruit 3' end processing factors is not present in mammalian Cajal bodies. It is also reasonable that factors involved in histone pre-mRNA processing would also be recruited to RNA Pol II when it is transcribing polyadenylated mRNAs. This raises the question of whether histone genes might signal for a novel modification of the CTD. Traditional dogma states that Ser2 and Ser5 are modified along genes, but recently phosphorylation of Ser7 has been identified as required for snRNA 3' end formation (Egloff et al., 2007b). The consensus repeat for the CTD is YSPTSPS, leaving open the possibility of modifications on Tyr1 or Thr4, or even a recruitment that does not involve the CTD. I demonstrate that SLBP is in a complex with RNA Pol II using both crosslinked and uncrosslinked IPs although the interaction might not be direct. The only antibody that coimmunoprecipitates SLBP is one which recognizes Rpb3, a smaller subunit of RNA Pol II, and therefore has no preference for CTD modification state; neither Ser2 nor Ser5 phosphospecific antibodies coimmunoprecipitate SLBP. This mechanism of recruitment is a topic which bears further studies and will, I believe, be critical in the regulation of histone gene expression.

### **Transcription termination of histone genes does not require Xrn2.**

Termination of transcription is tightly linked with 3' end formation and is essential for maintenance of transcriptome integrity, as failure to properly disengage RNA Pol II from the DNA template interferes with expression of downstream genes.

Currently there are two competing models, the allosteric and torpedo models (Lykke-Andersen and Jensen, 2007) which posit that an allosteric change in RNA Pol II or an exonuclease torpedo are responsible for termination respectively, though recent evidence implies that both models have a role (Luo et al., 2006a). For polyadenylated mammalian transcripts, the 5'-3' exonuclease, Xrn2, is recruited to genes via an interaction with CstF64 (Kaneko et al., 2007) and likely degrades the downstream cleavage fragment. However the CstF complex has no role in histone mRNA processing which begs the question; how is transcription of histone genes terminated?

In *Drosophila*, there are strong polymerase pause sites just downstream of the cleavage site (Adamson and Price, 2003) and RNA Pol II is not associated with the DNA in the intergenic region between histone H3 and H1 (Chapter two), implying that termination is a very efficient process. Since the *Drosophila* histone genes are located so close together, transcription must terminate between the genes to avoid production of double stranded RNA. Degradation of the downstream cleavage product in vitro is dependent on the U7 snRNP in both *Drosophila* and mammals (Dominski et al., 2005a), and is coupled with the processing reaction, consistent with a role for 3' end formation in transcription termination on histone genes. The endonuclease responsible for cleaving histone pre-mRNAs, CPSF73, has both endonucleolytic and 5'-3' exonucleolytic activities (Dominski et al., 2005b), allowing for the possibility that it may play a role in histone gene regulation in addition to cleavage.

Zbig Dominski and I hypothesized that CPSF73 is indeed the exonuclease which acts as the torpedo to terminate transcription of histone genes. I show in Chapter 3 that knockdown of Xrn2 does not affect the levels of downstream transcripts of a mammalian histone gene, but does significantly stabilize downstream RNA species from the GAPDH gene which encodes a poly(A) containing mRNA. Interestingly, knockdown of Xrn2 does have a mild stabilizing effect on RNA farther (~500 nts) downstream of the cleavage site indicating that CPSF73 may not be as processive an exonuclease as Xrn2. As degradation of the downstream cleavage product is U7 snRNP-dependent, should CPSF73 fall off of the transcript it is degrading, it would be unable to get back on. At this point an exonuclease such as Xrn2, which can initiate degradation de novo would have to complete degradation of the downstream transcript, and may account for the slight stabilization of the RNA fragments farther downstream.

### **SLBP functions in export of histone mRNA.**

SLBP is a multifunctional protein which is active in multiple aspects of histone mRNA metabolism [reviewed in (Marzluff et al., 2008)]. SLBP binds to the nascent RNA where it is essential for processing, in concert with the U7 snRNP. After a proper 3' end is formed, SLBP remains associated with the mature RNA and escorts it into the cytoplasm where it is active in translation. At the end of S-phase, SLBP participates in degradation of histone mRNAs before being phosphorylated by cyclinA/Cdk2 and degraded itself. Knockdown of SLBP by siRNA in mammalian cells results in an accumulation of cells in S-phase, but what is the cause of this accumulation? I asked which one of the functions of SLBP was responsible for this

phenotype. To test this, I created a panel of SLBP mutants each of which was resistant to the siRNA against SLBP; in this way I could remove endogenous SLBP, replace it with a mutant form lacking a known function and monitor rescue of the cell cycle phenotype. Surprisingly, all of the mutants tested including a truncated form of the protein containing only the RNA binding, RNA processing and translation domains, rescued the phenotype as well as the fully functional SLBP protein. I concluded that this assay was not sufficient to address the question of what the major lesion in histone mRNA metabolism is in cells severely depleted of SLBP and, along with Eric Wagner and Tom Mullen, set out to do a thorough molecular analysis of histone metabolism in SLBP deficient cells. I describe these results in Chapter 4 of this dissertation.

Tom Mullen determined that SLBP knockdown activates a DNA damage checkpoint likely due to a lack of histones available for deposition behind active replication forks. Interestingly, histone mRNA levels were only down about 2-fold in SLBP knockdown cells and the vast majority is properly processed. Tom and I discovered that there is polyadenylation of some of the RNA, but only in specific instances. There is some misprocessing of histone mRNAs in mammalian cells, as in *Drosophila*, but it is restricted to a few genes which have polyadenylation signals. When we cloned the polyadenylated RNAs we discovered that only a few genes generated polyadenylated transcripts. Intriguingly, some of these genes contained canonical poly(A) sites which overlapped with the histone pre-mRNA processing cis-elements, the stemloop or histone downstream element. These poly(A) sites would be masked by SLBP or the U7 snRNP under normal conditions of histone pre-mRNA

processing. Their conservation indicates a potential safety mechanism to allow for histone biosynthesis in the event of a misprocessing event. That they are used sparingly in mammals is not surprising as these RNAs are not cell cycle regulated and would result in histone overexpression if they accumulated to a significant extent. I also demonstrated that histone protein expression is down 2-3 fold in SLBP knockdown cells. The most surprising finding in these studies was that histone mRNA is retained in the nucleus in SLBP knockdown cells and that SLBP can promote export of a reporter. This identifies a new role for SLBP in nuclear export of histone mRNAs and helps rationalize a series of conflicting papers in the literature. In *Xenopus* oocytes, neither SLBP nor the SL is required for histone mRNA export, rather mRNA length is the determining factor. The mRNA must be sufficiently long for enough RNA binding proteins to assemble and recruit the TAP complex. This result likely reflects the high concentration of RNA binding proteins in the *Xenopus* germinal vesicle.

### **Final conclusions.**

Processing of pre-mRNA is a vital step in regulation of gene expression, and the replication dependent histone genes are no exception. These highly conserved genes have evolved their own mechanism of 3' end formation that is conserved from *Drosophila* to humans. Processing of histone pre-mRNAs, although unique, shares some components of the cleavage factor for canonical polyadenylated mRNAs. While assembly of the cleavage machinery at sites of processing is guided by different cis and trans elements for histone and poly(A) mRNAs, a core cleavage factor is common between the two.

Termination of transcription is another critical event for proper gene regulation. Termination of transcription for poly(A) mRNAs most likely involves two factors, Pcf11 and Xrn2, to disengage RNA Pol II from the DNA template. Here again, histone genes appear to have evolved a separate mechanism for this process. Termination of histone gene transcription is not dependent on Xrn2, rather using the exonuclease function of CPSF73 to achieve degradation of the downstream cleavage product and subsequent disengagement of RNA Pol II. This particular mechanism may be particularly important in lower eukaryotes, such as *Drosophila*, where the histone genes are located immediately adjacent to one another, and mistakes in transcription termination would generate double-stranded RNAs.

A common factor for many steps of histone mRNA metabolism is the histone-specific RNA binding protein SLBP. This protein binds the nascent transcript and is subsequently involved in cleavage, nuclear export, translation and degradation of histone mRNAs. SLBP embodies the nature of histone mRNA metabolism as it is unique to this class of mRNAs, yet facilitates interactions with the machinery used by canonical poly(A) mRNAs at all steps in the life of the RNA, yielding a unique mechanism of gene regulation.

## Reference List

1. Adamson, T.E., and Price, D.H. (2003). Cotranscriptional processing of *Drosophila* histone mRNAs. *Mol. Cell Biol.* 23, 4046-4055.
2. Adamson, T.E., Shutt, D.C., and Price, D.H. (2005). Functional coupling of cleavage and polyadenylation with transcription of mRNA. *J. Biol. Chem.* 280, 32262-32271.
3. Aravind, L. (1999). An evolutionary classification of the metallo-beta-lactamase fold proteins. *In Silico. Biol.* 1, 69-91.
4. Arnold, D.R., Francon, P., Zhang, J., Martin, K., and Clarke, H.J. (2008). Stem-loop binding protein expressed in growing oocytes is required for accumulation of mRNAs encoding histones H3 and H4 and for early embryonic development in the mouse. *Dev. Biol.* 313, 347-358.
5. Azzouz, T.N., Gruber, A., and Schumperli, D. (2005). U7 snRNP-specific Lsm11 protein: dual binding contacts with the 100 kDa zinc finger processing factor (ZFP100) and a ZFP100-independent function in histone RNA 3' end processing. *Nucleic Acids Res.* 33, 2106-2117.
6. Bai, Y., Auperin, T.C., Chou, C.Y., Chang, G.G., Manley, J.L., and Tong, L. (2007). Crystal structure of murine CstF-77: dimeric association and implications for polyadenylation of mRNA precursors. *Mol. Cell* 25, 863-875.
7. Barabino, S.M.L., Ohnacker, M., and Keller, M. (2000). Distinct roles of two Yth1p domains in 3'-end cleavage and polyadenylation of yeast pre-mRNAs. *EMBO J.* 19, 3778-3787.
8. Barcaroli, D., Bongiorno-Borbone, L., Terrinoni, A., Hofmann, T.G., Rossi, M., Knight, R.A., Matera, A.G., Melino, G., and De, L., V. (2006a). FLASH is required for histone transcription and S-phase progression. *Proc. Natl. Acad. Sci. U. S. A* 103, 14808-14812.
9. Barcaroli, D., Dinsdale, D., Neale, M.H., Bongiorno-Borbone, L., Ranalli, M., Munarriz, E., Sayan, A.E., McWilliam, J.M., Smith, T.M., Fava, E., Knight, R.A., Melino, G., and De, L., V. (2006b). FLASH is an essential component of Cajal bodies. *Proc. Natl. Acad. Sci. U. S. A* 103, 14802-14807.
10. Barilla, D., Lee, B.A., and Proudfoot, N.J. (2001). Cleavage/polyadenylation factor IA associates with the carboxyl-terminal domain of RNA

polymerase II in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U. S. A* 98, 445-450.

11. Benoit, B., Juge, F., Iral, F., Audibert, A., and Simonelig, M. (2002). Chimeric human CstF-77/*Drosophila* Suppressor of forked proteins rescue suppressor of forked mutant lethality and mRNA 3' end processing in *Drosophila*. *Proc. Natl. Acad. Sci. U. S. A* 99, 10593-10598.
12. Bentley, D.L. (2005). Rules of engagement: co-transcriptional recruitment of pre-mRNA processing factors. *Curr. Opin. Cell Biol.* 17, 251-256.
13. Bogenberger, J.M., and Laybourn, P.J. (2008). Human T Lymphotropic Virus Type 1 protein Tax reduces histone levels. *Retrovirology*. 5, 9-15.
14. Brown, K.M., and Gilmartin, G.M. (2003). A mechanism for the regulation of pre-mRNA 3' processing by human cleavage factor Im. *Mol. Cell* 12, 1467-1476.
15. Buratowski, S. (2005a). Connections between mRNA 3' end processing and transcription termination. *Curr. Opin. Cell Biol.* 17, 257-261.
16. Buratowski, S. (2005b). Connections between mRNA 3' end processing and transcription termination. *Curr. Opin. Cell Biol.* 17, 257-261.
17. Cakmakci, N.G., Lerner, R.S., Wagner, E.J., Zheng, L.-X., and Marzluff, W.F. (2008). SLIP1, A factor required for activation of histone mRNA translation by the stem-loop binding protein. *Mol. Cell Biol.* 28, 1182-1194.
18. Callebaut, I., Moshous, D., Mornon, J.P., and de Villartay, J.P. (2002). Metallo-beta-lactamase fold within nucleic acids processing enzymes: the beta-CASP family. *Nucleic Acids Res.* 30, 3592-3601.
19. Carmo-Fonseca, M., Pepperkok, R., Carvalho, M.T., and Lamond, A.I. (1992). Transcription-dependent colocalization of the U1, U2, U4/U6, and U5 snRNPs in coiled bodies. *J. Cell Biol.* 117, 1-14.
20. Chodchoy, N., Pandey, N.B., and Marzluff, W.F. (1991). An intact histone 3' processing site is required for transcription termination in a mouse histone H2a gene. *Mol. Cell. Biol.* 11, 497-509.
21. Cismowski, M.J., Laff, G.M., Solomon, M.J., and Reed, S.I. (1995). KIN28 encodes a C-terminal domain kinase that controls mRNA transcription in *Saccharomyces cerevisiae* but lacks cyclin-dependent kinase-activating kinase (CAK) activity. *Mol. Cell Biol.* 15, 2983-2992.
22. Colgan, D.F., and Manley, J.L. (1997). Mechanism and regulation of mRNA polyadenylation. *Genes and Development* 11, 2755-2766.



23. Coller, J.M., Gray, N.K., and Wickens, M.P. (1998). mRNA stabilization by poly(A) binding protein is independent of poly(A) and requires translation. *Genes and Development* 12, 3226-3235.
24. Cullen, B.R. (2004). Assaying nuclear messenger RNA export in human cells. *Methods Mol. Biol.* 257, 85-92.
25. de la Sierra-Gallay IL, Zig, L., Jamalli, A., and Putzer, H. (2008). Structural insights into the dual activity of RNase J. *Nat. Struct. Mol. Biol.* 15, 206-212.
26. Debry, R.W., and Marzluff, W.F. (1994). Selection on silent sites in the rodent H3 histone gene family. *Genetics* 138, 191-202.
27. DeLisle, A.J., Graves, R.A., Marzluff, W.F., and Johnson, L.F. (1983). Regulation of histone mRNA production and stability in serum stimulated mouse fibroblasts. *Mol. Cell. Biol.* 3, 1920-1929.
28. Dichtl, B., Blank, D., Sadowski, M., Hubner, W., Weiser, S., and Keller, W. (2002). Yhh1p/Cft1p directly links poly(A) site recognition and RNA polymerase II transcription termination. *EMBO J.* 21, 4125-4135.
29. Doherty, S.C., McKeown, S.R., Kelvey-Martin, V., Downes, C.S., Atala, A., Yoo, J.J., Simpson, D.A., and Kaufmann, W.K. (2003). Cell cycle checkpoint function in bladder cancer. *J. Natl. Cancer Inst.* 95, 1859-1868.
30. Dominski, Z. (2007). Nucleases of the metallo-beta-lactamase family and their role in DNA and RNA metabolism. *Crit Rev. Biochem. Mol. Biol.* 42, 67-93.
31. Dominski, Z., Erkmann, J.A., Greenland, J.A., and Marzluff, W.F. (2001). Mutations in the RNA binding domain of stem-loop binding protein define separable requirements for RNA binding and histone pre-mRNA processing. *Mol. Cell. Biol.* 21, 2008-2017.
32. Dominski, Z., Erkmann, J.A., Yang, X., Sanchez, R., and Marzluff, W.F. (2002). A novel zinc finger protein is associated with U7 snRNP and interacts with the stem-loop binding protein in the histone pre-mRNP to stimulate 3'-end processing. *Genes Dev.* 16, 58-71.
33. Dominski, Z., and Marzluff, W.F. (2007). Formation of the 3' end of histone mRNA: Getting closer to the end. *Gene* 396, 373-390.
34. Dominski, Z., Yang, X., Purdy, M., and Marzluff, W.F. (2005a). Differences and similarities between *Drosophila* and mammalian 3' end processing of histone pre-mRNAs. *RNA* 11, 1835-1847.

35. Dominski, Z., Yang, X.C., and Marzluff, W.F. (2005b). The polyadenylation factor CPSF73 is involved in histone pre-mRNA processing. *Cell* 123, 37-48.
36. Dominski, Z., Zheng, L.-X., Sanchez, R., and Marzluff, W.F. (1999). The stem-loop binding protein facilitates 3' end formation by stabilizing U7 snRNP binding to the histone pre-mRNA. *Mol. Cell. Biol.* 19, 3561-3570.
37. Eckner, R., Ellmeier, W., and Birnstiel, M.L. (1991). Mature mRNA 3' end formation stimulates RNA export from the nucleus. *EMBO J.* 10, 3513-3522.
38. Egloff, S., O'Reilly, D., Chapman, R.D., Taylor, A., Tanzhaus, K., Pitts, L., Eick, D., and Murphy, S. (2007a). Serine-7 of the RNA polymerase II CTD is specifically required for snRNA gene expression. *Science* 318, 1777-1779.
39. Egloff, S., O'Reilly, D., Chapman, R.D., Taylor, A., Tanzhaus, K., Pitts, L., Eick, D., and Murphy, S. (2007b). Serine-7 of the RNA polymerase II CTD is specifically required for snRNA gene expression. *Science* 318, 1777-1779.
40. Eitoku, M., Sato, L., Senda, T., and Horikoshi, M. (2008). Histone chaperones: 30 years from isolation to elucidation of the mechanisms of nucleosome assembly and disassembly. *Cell Mol. Life Sci.* 65, 414-444.
41. Erkmann, J.A., and Kutay, U. (2004). Nuclear export of mRNA: from the site of transcription to the cytoplasm. *Exp. Cell Res.* 296, 12-20.
42. Erkmann, J.A., Sanchez, R., Treichel, N., Marzluff, W.F., and Kutay, U. (2005a). Nuclear export of metazoan replication-dependent histone mRNAs is dependent on RNA length and is mediated by TAP. *RNA* 11, 45-58.
43. Erkmann, J.A., Wagner, E.J., Dong, J., Zhang, Y.P., Kutay, U., and Marzluff, W.F. (2005b). Nuclear import of the stem-loop binding protein and localization during the cell cycle. *Mol. Biol. Cell* 16, 2960-2971.
44. Even, S., Pellegrini, O., Zig, L., Labas, V., Vinh, J., Brechemmier-Baey, D., and Putzer, H. (2005). Ribonucleases J1 and J2: two novel endoribonucleases in *B.subtilis* with functional homology to *E.coli* RNase E. *Nucleic Acids Res.* 33, 2141-2152.
45. Friend, K., Lovejoy, A.F., and Steitz, J.A. (2007). U2 snRNP binds intronless histone pre-mRNAs to facilitate U7 snRNP dependent 3' end formation. *Molecular Cell* 28, 240-252.

46. Ghule, P.N., Dominski, Z., Yang, X.C., Marzluff, W.F., Becker, K.A., Harper, J.W., Lian, J.B., Stein, J.L., Van Wijnen, A.J., and Stein, G.S. (2008). Staged assembly of histone gene expression machinery at subnuclear foci in the abbreviated cell cycle of human embryonic stem cells. *Proc. Natl. Acad. Sci. U. S. A* 44, 16964-16969.
47. Gick, O., Krämer, A., Vasserot, A., and Birnstiel, M.L. (1987). Heat-labile regulatory factor is required for 3' processing of histone precursor mRNAs. *Proc. Natl. Acad. Sci. USA* 84, 8937-8940.
48. Gilchrist, D.A., Nechaev, S., Lee, C., Ghosh, S.K., Collins, J.B., Li, L., Gilmour, D.S., and Adelman, K. (2008). NELF-mediated stalling of Pol II can enhance gene expression by blocking promoter-proximal nucleosome assembly. *Genes Dev.* 22, 1921-1933.
49. Glover-Cutter, K., Kim, S., Espinosa, J., and Bentley, D.L. (2008). RNA polymerase II pauses and associates with pre-mRNA processing factors at both ends of genes. *Nat. Struct. Mol. Biol.* 15, 71-78.
50. Godfrey, A.C., Kupsco, J.M., Burch, B.D., Zimmerman, R.M., Dominski, Z., Marzluff, W.F., and Duronio, R.J. (2006). U7 snRNA mutations in *Drosophila* block histone pre-mRNA processing and block oogenesis. *RNA* 12, 396-409.
51. Gorgoni, B., Andrews, S., Schaller, A., Schumperli, D., Gray, N.K., and Muller, B. (2005). The stem-loop binding protein stimulates histone translation at an early step in the initiation pathway. *RNA* 11, 1030-1042.
52. Graves, R.A., Pandey, N.B., Chodchoy, N., and Marzluff, W.F. (1987). Translation is required for regulation of histone mRNA degradation. *Cell* 48, 615-626.
53. Gunjan, A., and Verreault, A. (2003). A Rad53 kinase-dependent surveillance mechanism that regulates histone protein levels in *S. cerevisiae*. *Cell* 115, 537-549.
54. Hatton, L.S., Eloranta, J.J., Figueiredo, L.M., Takagaki, Y., Manley, J.L., and O'Hare, K. (2000). The *Drosophila* homologue of the 64 kDa subunit of cleavage stimulation factor interacts with the 77 kDa subunit encoded by the suppressor of forked gene. *Nucleic Acids Res.* 28, 520-526.
55. Helmling, S., Zhelkovsky, A., and Moore, C.L. (2001). Fip1 regulates the activity of Poly(A) polymerase through multiple interactions. *Mol. Cell Biol.* 21, 2026-2037.
56. Hereford, L., Fahrner, K., Woolford, J.J., and Rosbash, M. (1979). Isolation of yeast histone genes H2a and H2b. *Cell* 18, 1261-1271.

57. Hirose, Y., and Manley, J.L. (1998). RNA polymerase II is an essential mRNA polyadenylation factor. *Nature* 395, 93-96.
58. Ho, C.K., and Shuman, S. (1999). Distinct roles for CTD Ser-2 and Ser-5 phosphorylation in the recruitment and allosteric activation of mammalian mRNA capping enzyme. *Mol. Cell* 3, 405-411.
59. Hoek, M., and Stillman, B. (2003). Chromatin assembly factor 1 is essential and couples chromatin assembly to DNA replication in vivo. *Proc. Natl. Acad. Sci. U. S. A* 100, 12183-12188.
60. Hollingworth, D., Noble, C.G., Taylor, I.A., and Ramos, A. (2006). RNA polymerase II CTD phosphopeptides compete with RNA for the interaction with Pcf11. *RNA*. 12, 555-560.
61. Huang, Y.Q., and Carmichael, G.G. (1996). Role of polyadenylation in nucleocytoplasmic transport of mRNA. *Mol. Cell. Biol.* 16, 1534-1542.
62. Huang, Y.Q., and Carmichael, G.G. (1997). The mouse histone H2a gene contains a small element that facilitates cytoplasmic accumulation of intronless gene transcripts and of unspliced HIV-1-related mRNAs. *Proc. Natl. Acad. Sci. USA* 94, 10104-10109.
63. Huang, Y.Q., and Steitz, J.A. (2001). Splicing factors SRp20 and 9G8 promote the nucleocytoplasmic export of mRNA. *Molecular Cell* 7, 899-905.
64. Jones, J.C., Phatnani, H.P., Haystead, T.A., MacDonald, J.A., Alam, S.M., and Greenleaf, A.L. (2004). C-terminal repeat domain kinase I phosphorylates Ser2 and Ser5 of RNA polymerase II C-terminal domain repeats. *J. Biol. Chem.* 279, 24957-24964.
65. Juge, F., Audibert, A., Benoit, B., and Simonelig, M. (2000). Tissue-specific autoregulation of *Drosophila* suppressor of forked by alternative poly(A) site utilization leads to accumulation of the suppressor of forked protein in mitotically active cells. *RNA*. 6, 1529-1538.
66. Kaneko, S., Rozenblatt-Rosen, O., Meyerson, M., and Manley, J.L. (2007). The multifunctional protein p54nrb/PSF recruits the exonuclease XRN2 to facilitate pre-mRNA 3' processing and transcription termination. *Genes Dev.* 21, 1779-1789.
67. Kaufmann, I., Martin, G., Friedlein, A., Langen, H., and Keller, W. (2004). Human Fip1 is a subunit of CPSF that binds to U-rich RNA elements and stimulates poly(A) polymerase. *EMBO J.* 23, 616-626.

68. Kaygun, H., and Marzluff, W.F. (2005a). Regulated degradation of replication-dependent histone mRNAs requires both ATR and Upf1. *Nat. Struct. Mol. Biol.* 12, 794-800.
69. Kaygun, H., and Marzluff, W.F. (2005b). Translation termination is involved in histone mRNA degradation when DNA replication is inhibited. *Mol. Cell Biol.* 25, 6879-6888.
70. Keller, W., Bienroth, S., Lang, K.M., and Christofori, G. (1991). Cleavage and polyadenylation factor CPF specifically interacts with the pre-mRNA 3' processing signal AAUAAA. *EMBO J.* 10, 4241-4249.
71. Keon, B.H., Schafer, S., Kuhn, C., Grund, C., and Franke, W.W. (1996). Symplekin, a novel type of tight junction plaque protein. *J. Cell Biol.* 134, 1003-1018.
72. Kim, M., Krogan, N.J., Vasiljeva, L., Rando, O.J., Nedeia, E., Greenblatt, J.F., and Buratowski, S. (2004). The yeast Rat1 exonuclease promotes transcription termination by RNA polymerase II. *Nature* 432, 517-522.
73. Kolev, N.G., and Steitz, J.A. (2005). Symplekin and multiple other polyadenylation factors participate in 3'-end maturation of histone mRNAs. *Genes Dev.* 19, 2583-2592.
74. Kolev, N.G., Yario, T.A., Benson, E., and Steitz, J.A. (2008). Conserved motifs in both CPSF73 and CPSF100 are required to assemble the active endonuclease for histone mRNA 3'-end maturation. *EMBO Rep.*
75. Komarnitsky, P., Cho, E.J., and Buratowski, S. (2000). Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. *Genes and Development* 14, 2452-2460.
76. Kornblihtt, A.R., de la, M.M., Fededa, J.P., Munoz, M.J., and Nogues, G. (2004). Multiple links between transcription and splicing. *RNA.* 10, 1489-1498.
77. Koseoglu, M.M., Graves, L.M., and Marzluff, W.F. (2008). Phosphorylation of threonine 61 by cyclin A/Cdk1 triggers degradation of stemloop binding protein at the end of S-phase. *Molecular & Cellular Biology* 28, 4469-4479.
78. Kyburz, A., Sadowski, M., Dichtl, B., and Keller, W. (2003). The role of the yeast cleavage and polyadenylation factor subunit Ydh1p/Cft2p in pre-mRNA 3'-end formation. *Nucleic Acids Res.* 31, 3936-3945.
79. Lanzotti, D.J., Kaygun, H., Yang, X., Duronio, R.J., and Marzluff, W.F. (2002). Developmental control of histone mRNA and dSLBP synthesis during

*Drosophila* embryogenesis and the role of dSLBP in histone mRNA 3' processing in vivo. *Mol. Cell Biol.* 22, 2267-2282.

80. Li, X., and Manley, J.L. (2005). Inactivation of the SR protein splicing factor ASF/SF2 results in genomic instability. *Cell* 122, 365-378.
81. Licatalosi, D.D., Geiger, G., Minet, M., Schroeder, S., Cilli, K., McNeil, J.B., and Bentley, D.L. (2002a). Functional interaction of yeast pre-mRNA 3' end processing factors with RNA polymerase II. *Molecular Cell* 9, 1101-1111.
82. Licatalosi, D.D., Geiger, G., Minet, M., Schroeder, S., Cilli, K., McNeil, J.B., and Bentley, D.L. (2002b). Functional interaction of yeast pre-mRNA 3' end processing factors with RNA polymerase II. *Mol. Cell* 9, 1101-1111.
83. Lifton, R.P., Goldberg, M.L., Karp, R.W., and Hogness, D.S. (1978). The organization of the histone genes in *drosophila melanogaster*: functional and evolutionary implications. *Cold Spring Harbor Symp. Quant. Biol.* 42, 1047-1051.
84. Listerman, I., Sapra, A.K., and Neugebauer, K.M. (2006). Cotranscriptional coupling of splicing factor recruitment and precursor messenger RNA splicing in mammalian cells. *Nat. Struct. Mol. Biol.* 13, 815-822.
85. Liu, J.L., Murphy, C., Buszczak, M., Clatterbuck, S., Goodman, R., and Gall, J.G. (2006). The *Drosophila melanogaster* Cajal body. *J. Cell Biol.* 172, 875-884.
86. Liu, T.-J., Levine, B.J., Skoultschi, A.I., and Marzluff, W.F. (1989). The efficiency of 3'-end formation contributes to the relative levels of different histone mRNAs. *Mol. Cell. Biol.* 9, 3499-3508.
87. Luger, K., Mäder, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8Å resolution. *Nature* 389, 251-260.
88. Luo, W., and Bentley, D. (2004). A ribonucleolytic rat torpedo does RNA polymerase II. *Cell* 119, 911-914.
89. Luo, W., Johnson, A.W., and Bentley, D.L. (2006a). The role of Rat1 in coupling mRNA 3'-end processing to transcription termination: implications for a unified allosteric-torpedo model. *Genes Dev.* 20, 954-965.
90. Luo, W., Johnson, A.W., and Bentley, D.L. (2006b). The role of Rat1 in coupling mRNA 3'-end processing to transcription termination:

- implications for a unified allosteric-torpedo model. *Genes Dev.* 20, 954-965.
91. Lykke-Andersen, S., and Jensen, T.H. (2007). Overlapping pathways dictate termination of RNA polymerase II transcription. *Biochimie* 89, 1177-1182.
  92. Ma, T.L., Van Tine, B.A., Wei, Y., Garrett, M.D., Nelson, D., Adams, P.D., Wang, J., Qin, J., Chow, L.T., and Harper, J.W. (2000). Cell cycle-regulated phosphorylation of p220<sup>NPAT</sup> by cyclin E/Cdk2 in Cajal bodies promotes histone gene transcription. *Genes and Development* 14, 2298-2313.
  93. Mandel, C.R., Bai, Y., and Tong, L. (2007). Protein factors in pre-mRNA 3'-end processing. *Cell Mol. Life Sci.* 65, 1099-1122.
  94. Mandel, C.R., Bai, Y., and Tong, L. (2008). Protein factors in pre-mRNA 3'-end processing. *Cell Mol. Life Sci.* 65, 1099-1122.
  95. Mandel, C.R., Kaneko, S., Zhang, H., Gebauer, D., Vethantham, V., Manley, J.L., and Tong, L. (2006). Polyadenylation factor CPSF73 is the pre-mRNA 3'-end-processing endonuclease. *Nature* 444, 953-956.
  96. Marshall, N.F., Peng, J.M., Xie, Z., and Price, D.H. (1996). Control of RNA polymerase II elongation potential by a novel carboxyl-terminal domain kinase. *J. Biol. Chem.* 271, 27176-27183.
  97. Marzluff, W.F. (2005). Metazoan replication dependent histone mRNAs: a unique class of RNA polymerase II transcripts. *Current Opinion in Cell Biology* 17, 274-280.
  98. Marzluff, W.F., Gongidi, P., Woods, K.R., Jin, J.P., and Maltais, L. (2002). The human and mouse replication-dependent histone genes. *Genomics* 80, 487-498.
  99. Marzluff, W.F., Wagner, E.J., and Duronio, R.J. (2008). Metabolism and regulation of canonical histone mRNAs: life without a poly(A) tail. *Nat. Rev. Genet.* 9, 843-854.
  100. Masuyama, K., Taniguchi, I., Kataoka, N., and Ohno, M. (2004). RNA length defines RNA export pathway. *Genes Dev.* 18, 2074-2085.
  101. Mathy, N., Benard, L., Pellegrini, O., Daou, R., Wen, T., and Condon, C. (2007a). 5'-to-3' exoribonuclease activity in bacteria: role of RNase J1 in rRNA maturation and 5' stability of mRNA. *Cell* 129, 681-692.

102. Mathy, N., Benard, L., Pellegrini, O., Daou, R., Wen, T., and Condon, C. (2007b). 5'-to-3' exoribonuclease activity in bacteria: role of RNase J1 in rRNA maturation and 5' stability of mRNA. *Cell* 129, 681-692.
103. McCracken, S., Fong, N., Yankulov, K., Ballantyne, S., Pan, G.H., Greenblatt, J., Patterson, S.D., Wickens, M., and Bentley, D.L. (1997). The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. *Nature* 385, 357-361.
104. Meeks-Wagner, D., and Hartwell, L.H. (1986). Normal stoichiometry of histone dimer sets is necessary for high fidelity of mitotic chromosome transmission. *Cell* 44, 43-52.
105. Misteli, T., and Spector, D.L. (1999). RNA polymerase II targets pre-mRNA splicing factors to transcription sites in vivo. *Mol. Cell* 3, 697-705.
106. Moore, M.J., Schwartzfarb, E.M., Silver, P.A., and Yu, M.C. (2006). Differential recruitment of the splicing machinery during transcription predicts genome-wide patterns of mRNA splicing. *Mol. Cell* 24, 903-915.
107. Morris, D.P., and Greenleaf, A.L. (2000). The splicing factor, Prp40, binds the phosphorylated carboxyl-terminal domain of RNA polymerase II. *J. Biol. Chem.* 275, 39935-39943.
108. Moss, S.B., Ferry, R.A., and Groudine, M. (1994). An alternative pathway of histone mRNA 3' end formation in mouse round spermatids. *Nucleic Acids Res.* 22, 3160-3166.
109. Mowry, K.L., Oh, R., and Steitz, J.A. (1989). Each of the conserved sequence elements flanking the cleavage site of mammalian histone pre-mRNAs has a distinct role in the 3'-end processing reaction. *Mol. Cell. Biol.* 9, 3105-3108.
110. Mowry, K.L., and Steitz, J.A. (1987). Identification of the human U7 snRNP as one of several factors involved in the 3' end maturation of histone premessenger RNA's. *Science* 238, 1682-1687.
111. Mullen, T.E., and Marzluff, W.F. (2008). Degradation of histone mRNA requires oligouridylation followed by decapping and simultaneous degradation of the mRNA both 5' to 3' and 3' to 5'. *Genes Dev.* 22, 50-65.
112. Murthy, K.G., and Manley, J.L. (1992). Characterization of the multisubunit cleavage-polyadenylation specificity factor from calf thymus. *J. Biol. Chem.* 267, 14804-14811.



113. Murthy, K.G.K., and Manley, J.L. (1995). The 160-kD subunit of human cleavage polyadenylation specificity factor coordinates pre-mRNA 3'-end formation. *Genes and Development* 9, 2672-2683.
114. Narita, T., Yung, T.M., Yamamoto, J., Tsuboi, Y., Tanabe, H., Tanaka, K., Yamaguchi, Y., and Handa, H. (2007). NELF Interacts with CBC and Participates in 3' End Processing of Replication-Dependent Histone mRNAs. *Mol. Cell* 26, 349-365.
115. Neugebauer, K.M. (2002). On the importance of being co-transcriptional. *J. Cell Sci.* 115, 3865-3871.
116. Ni, Z., Saunders, A., Fuda, N.J., Yao, J., Suarez, J.R., Webb, W.W., and Lis, J.T. (2008). P-TEFb is critical for the maturation of RNA polymerase II into productive elongation in vivo. *Mol. Cell Biol.* 28, 1161-1170.
117. O'Connor, J.P., and Peebles, C.L. (1992a). PTA1, an essential gene of *Saccharomyces cerevisiae* affecting pre-tRNA processing. *Mol. Cell Biol.* 12, 3843-3856.
118. O'Connor, J.P., and Peebles, C.L. (1992b). PTA1, an essential gene of *Saccharomyces cerevisiae* affecting pre-tRNA processing. *Mol. Cell Biol.* 12, 3843-3856.
119. Pandey, N.B., Chodchoy, N., Liu, T.-J., and Marzluff, W.F. (1990). Introns in histone genes alter the distribution of 3' ends. *Nucleic Acids Res.* 18, 3161-3170.
120. Pandey, N.B., and Marzluff, W.F. (1987). The stem-loop structure at the 3' end of histone mRNA is necessary and sufficient for regulation of histone mRNA stability. *Mol. Cell. Biol.* 7, 4557-4559.
121. Pandey, N.B., Williams, A.S., Sun, J.-H., Brown, V.D., Bond, U., and Marzluff, W.F. (1994). Point mutations in the stem-loop at the 3' end of mouse histone mRNA reduce expression by reducing the efficiency of 3' end formation. *Mol. Cell. Biol.* 14, 1709-1720.
122. Pardue, M.L., Kedes, L.H., Weinberg, E.S., and Birnstiel, M.L. (1977). Localization of sequences coding for histone messenger RNA in the chromosomes of *Drosophila melanogaster*. *Chromosoma* 63, 135-151.
123. Parker, R., and Song, H. (2004). The enzymes and control of eukaryotic mRNA turnover. *Nat. Struct. Mol. Biol.* 11, 121-127.
124. Phatnani, H.P., and Greenleaf, A.L. (2004). Identifying PhosphoCTD-Associating Proteins. *Methods Mol. Biol.* 257, 17-28.

125. Phatnani, H.P., and Greenleaf, A.L. (2006). Phosphorylation and functions of the RNA polymerase II CTD. *Genes Dev.* 20, 2922-2936.
126. Pilch, D.R., and Marzluff, W.F. (1991). Expression of histone-U1 snRNA genes: U1 promoters are compatible with histone 3' end formation. *Gene Expression* 1, 41-53.
127. Pillai, R.S., Grimmier, M., Meister, G., Will, C.L., Luhrmann, R., Fischer, U., and Schumperli, D. (2003). Unique Sm core structure of U7 snRNPs: assembly by a specialized SMN complex and the role of a new component, Lsm11, in histone RNA processing. *Genes Dev.* 17, 2321-2333.
128. Pillai, R.S., Will, C.L., Lührmann, R., Schümperli, D., and Müller, B. (2001). Purified U7 snRNPs lack the Sm proteins D1 and D2 but contain Lsm10, a new 14 kDa Sm D1-like protein. *EMBO J.* 20, 5470-5479.
129. Preiss, T., and Hentze, M.W. (1998). Dual function of the messenger RNA cap structure in poly(A)-tail-promoted translation in yeast. *Nature* 392, 516-520.
130. Preker, P.J., Lingner, J., Minvielle-Sebastia, L., and Keller, W. (1995). The FIP1 gene encodes a component of a yeast pre-mRNA polyadenylation factor that directly interacts with poly(A) polymerase. *Cell* 81, 379-389.
131. Preker, P.J., Ohnacker, M., Minvielle-Sebastia, L., and Keller, W. (1997). A multisubunit 3' end processing factor from yeast containing poly(A) polymerase and homologues of the subunits of mammalian cleavage and polyadenylation specificity factor. *EMBO J.* 16, 4727-4737.
132. Proudfoot, N. (2004). New perspectives on connecting messenger RNA 3' end formation to transcription. *Curr. Opin. Cell Biol.* 16, 272-278.
133. Proudfoot, N.J., Furger, A., and Dye, M.J. (2002). Integrating rRNA processing with transcription. *Cell* 108, 501-512.
134. Rosonina, E., Kaneko, S., and Manley, J.L. (2006). Terminating the transcript: breaking up is hard to do. *Genes Dev.* 20, 1050-1056.
135. Ruegsegger, U., Blank, D., and Keller, W. (1998). Human pre-mRNA cleavage factor Im is related to spliceosomal SR proteins and can be reconstituted in vitro from recombinant subunits. *Molecular Cell* 1, 243-253.
136. Ryan, K., Calvo, O., and Manley, J.L. (2004). Evidence that polyadenylation factor CPSF73 is the mRNA 3' processing endonuclease. *RNA.* 10, 565-573.

137. Sadowski, M., Dichtl, B., Hubner, W., and Keller, W. (2003). Independent functions of yeast Pcf11p in pre-mRNA 3' end processing and in transcription termination. *EMBO J.* 22, 2167-2177.
138. Sancar, A., Lindsey-Boltz, L.A., Unsal-Kacmaz, K., and Linn, S. (2004). Molecular mechanisms of mammalian DNA repair and the DNA damage checkpoints. *Annu. Rev. Biochem.* 73, 39-85.
139. Sanchez, R., and Marzluff, W.F. (2002). The stem-loop binding protein is required for efficient translation of histone mRNA in vivo and in vitro. *Mol. Cell. Biol.* 22, 7093-7104.
140. Sarma, K., and Reinberg, D. (2005). Histone variants meet their match. *Nat. Rev. Mol. Cell Biol.* 6, 139-149.
141. Scharl, E.C., and Steitz, J.A. (1994). The site of 3' end formation of histone messenger RNA is a fixed distance from the downstream element recognized by the U7 snRNP. *EMBO J.* 13, 2432-2440.
142. Schochetman, G., and Perry, R.P. (1972). Early appearance of histone messenger RNA in polyribosomes of cultured L cells. *J. Mol. Biol.* 63, 591-596.
143. Schroeder, S.C., Schwer, B., Shuman, S., and Bentley, D. (2000). Dynamic association of capping enzymes with transcribing RNA polymerase II. *Genes and Development* 14, 2435-2440.
144. Shi, Y., Di Giammartino, D.C., Taylor, D., Sarkeshik, A., Rice, W.J., Yates, J.R., III, Frank, J., and Manley, J.L. (2009). Molecular architecture of the human pre-mRNA 3' processing complex. *Mol. Cell* 33, 365-376.
145. Shibahara, K., and Stillman, B. (1999). Replication-dependent marking of DNA by PCNA facilitates CAF-1-coupled inheritance of chromatin. *Cell* 96, 575-585.
146. Shopland, L.S., Byron, M., Stein, J.L., Lian, J.B., Stein, G.S., and Lawrence, J.B. (2001). Replication-dependent histone gene expression is related to Cajal body (CB) association but does not require sustained CB contact. *Molecular Biology of the Cell* 12, 565-576.
147. Skaar, D.A., and Greenleaf, A.L. (2002). The RNA polymerase II CTD kinase CTDK-I affects pre-mRNA 3' cleavage/polyadenylation through the processing component Pti1p. *Mol. Cell* 10, 1429-1439.
148. Solinger, J.A., Pascolini, D., and Heyer, W.D. (1999). Active-site mutations in the Xrn1p exoribonuclease of *Saccharomyces cerevisiae* reveal a specific role in meiosis. *Mol. Cell. Biol.* 19, 5930-5942.

149. Stevens, A. (1978). An exoribonuclease from *Saccharomyces cerevisiae*: effect of modifications of 5' end groups on the hydrolysis of substrates to 5' mononucleotides. *Biochem. Biophys. Res. Commun.* 81, 656-661.
150. Stevens, A. (1980). Purification and characterization of a *Saccharomyces cerevisiae* exoribonuclease which yields 5'-mononucleotides by a 5' leads to 3' mode of hydrolysis. *J. Biol. Chem.* 255, 3080-3085.
151. Stevens, A., and Poole, T.L. (1995). 5'-exonuclease-2 of *Saccharomyces cerevisiae*. Purification and features of ribonuclease activity with comparison to 5'-exonuclease-1. *J. Biol. Chem.* 270, 16063-16069.
152. Strahl, B.D., and Allis, C.D. (2000). The language of covalent histone modifications. *Nature* 403, 41-45.
153. Sullivan, E., Santiago, C., Parker, E.D., Dominski, Z., Yang, X., Lanzotti, D.J., Ingledue, T.C., Marzluff, W.F., and Duronio, R.J. (2001). *Drosophila* stem loop binding protein coordinates accumulation of mature histone mRNA with cell cycle progression. *Genes Dev.* 15, 173-187.
154. Sullivan, K.D., Mullen, T.E., Marzluff, W.F., and Wagner, E.J. (2009). Knockdown of SLBP results in nuclear retention of histone mRNA. ()
155. Swinburne, I.A., Meyer, C.A., Liu, X.S., Silver, P.A., and Brodsky, A.S. (2006). Genomic localization of RNA binding proteins reveals links between pre-mRNA processing and transcription. *Genome Res.* 16, 912-921.
156. Takagaki, Y., and Manley, J.L. (1997). RNA recognition by the human polyadenylation factor CstF. *Mol. Cell Biol.* 17, 3907-3914.
157. Takagaki, Y., and Manley, J.L. (2000a). Complex protein interactions within the human polyadenylation machinery identify a novel component. *Mol. Cell Biol.* 20, 1515-1525.
158. Takagaki, Y., and Manley, J.L. (2000b). Complex protein interactions within the human polyadenylation machinery identify a novel component. *Mol. Cell Biol.* 20, 1515-1525.
159. Takagaki, Y., Manley, J.L., MacDonald, C.C., Wilusz, J., and Shenk, T. (1990). A multisubunit factor, CstF, is required for polyadenylation of mammalian pre-mRNAs. *Genes Dev.* 4, 2112-2120.
160. Taube, R., Lin, X., Irwin, D., Fujinaga, K., and Peterlin, B.M. (2002). Interaction between P-TEFb and the C-terminal domain of RNA polymerase II activates transcriptional elongation from sites upstream or downstream of target genes. *Mol. Cell Biol.* 22, 321-331.

161. Tollervey, D. (2004). Molecular biology: termination by torpedo. *Nature* 432, 456-457.
162. Townley-Tilson, W.H., Pendergrass, S.A., Marzluff, W.F., and Whitfield, M.L. (2006). Genome-wide analysis of mRNAs bound to the histone stemloop binding protein. *RNA* 12, 1853-1867.
163. Tyler, J.K., Adams, C.R., Chen, S.R., Kobayashi, R., Kamakaka, R.T., and Kadonaga, J.T. (1999). The RCAF complex mediates chromatin assembly during DNA replication and repair. *Nature* 402, 555-560.
164. Venkataraman, K., Brown, K.M., and Gilmartin, G.M. (2005). Analysis of a noncanonical poly(A) site reveals a tripartite mechanism for vertebrate poly(A) site recognition. *Genes Dev.* 19, 1315-1327.
165. Vethantham, V., Rao, N., and Manley, J.L. (2007). Sumoylation modulates the assembly and activity of the pre-mRNA 3' processing complex. *Mol. Cell Biol.* 27, 8848-8858.
166. Wagner, E.J., and Marzluff, W.F. (2006). ZFP100, a component of the active U7 snRNP limiting for histone pre-mRNA processing, is required for entry into S-phase. *Molecular & Cellular Biology* 26, 6702-6712.
167. Wagner, E.J., Berkow, A., and Marzluff, W.F. (2005). Expression of an RNAi-resistant SLBP restores proper S-phase progression. *Biochem. Soc. Trans.* 33, 471-473.
168. Wagner, E.J., Burch, B.D., Godfrey, A.C., Salzer, H.R., Duronio, R.J., and Marzluff, W.F. (2007). A genome-wide RNA interference screen reveals that variant histones are necessary for replication-dependent histone pre-mRNA processing. *Molecular Cell* 28, 692-699.
169. Wagner, E.J., and Garcia-Blanco, M.A. (2002). RNAi-Mediated PTB Depletion Leads to Enhanced Exon Definition. *Mol. Cell* 10, 943-949.
170. Wagner, E.J., Ospina, J.K., Hu, Y., Dundr, M., Matera, A.G., and Marzluff, W.F. (2006). Conserved zinc fingers mediate multiple function of ZFP100, a U7 snRNP associated protein. *RNA* 12, 1206-1218.
171. Walther, T.N., Wittop, K.T., Schümperli, D., and Muller, B. (1998). A 5'-3' exonuclease activity involved in forming the 3' products of histone pre-mRNA processing in vitro. *RNA* 4, 1034-1046.
172. Wang, Z.-F., Whitfield, M.L., Ingledue, T.I., Dominski, Z., and Marzluff, W.F. (1996). The protein which binds the 3' end of histone mRNA: a novel RNA-binding protein required for histone pre-mRNA processing. *Genes and Development* 10, 3028-3040.

173. Weitzer, S., and Martinez, J. (2007). The human RNA kinase hClp1 is active on 3' transfer RNA exons and short interfering RNAs. *Nature* 447, 222-226.
174. West, S., Gromak, N., and Proudfoot, N.J. (2004). Human 5' → 3' exonuclease Xrn2 promotes transcription termination at co-transcriptional cleavage sites. *Nature* 432, 522-525.
175. White, A.E., Leslie, M.E., Calvi, B.R., Marzluff, W.F., and Duronio, R.J. (2007). Developmental and cell cycle regulation of the *Drosophila* histone locus body. *Mol. Biol. Cell* 18, 2491-2502.
176. Whitelaw, E., Coates, A., and Proudfoot, N.J. (1986). Globin gene transcripts can utilize histone gene 3' end processing signals. *Nucleic Acids Res.* 14, 7059-7070.
177. Whitfield, M.L., Kaygun, H., Erkmann, J.A., Townley-Tilson, W.H.D., Dominski, Z., and Marzluff, W.F. (2004). SLBP is associated with histone mRNA on polyribosomes as a component of histone mRNP. *Nucleic Acids Res.* 32, 4833-4842.
178. Whitfield, M.L., Zheng, L.-X., Baldwin, A., Ohta, T., Hurt, M.M., and Marzluff, W.F. (2000). Stem-loop binding protein, the protein that binds the 3' end of histone mRNA, is cell cycle regulated by both translational and posttranslational mechanisms. *Mol. Cell Biol.* 20, 4188-4198.
179. Wiegand, H.L., Coburn, G.A., Zeng, Y., Kang, Y., Bogerd, H.P., and Cullen, B.R. (2002). Formation of Tap/NXT1 heterodimers activates Tap-dependent nuclear mRNA export by enhancing recruitment to nuclear pore complexes. *Mol. Cell Biol.* 22, 245-256.
180. Xiao, T., Shibata, Y., Rao, B., Laribee, R.N., O'Rourke, R., Buck, M.J., Greenblatt, J.F., Krogan, N.J., Lieb, J.D., and Strahl, B.D. (2007). The RNA polymerase II kinase Ctk1 regulates positioning of a 5' histone methylation boundary along genes. *Mol. Cell Biol.* 27, 721-731.
181. Xie, S.Q., and Pombo, A. (2006). Distribution of different phosphorylated forms of RNA polymerase II in relation to Cajal and PML bodies in human cells: an ultrastructural study. *Histochem. Cell Biol.* 125, 21-31.
182. Yang, J., Bogerd, H.P., Wang, P.J., Page, D.C., and Cullen, B.R. (2001). Two closely related human nuclear export factors utilize entirely distinct export pathways. *Molecular Cell* 8, 397-406.
183. Yang, X.C., Sullivan, K.D., Marzluff, W.F., and Dominski, Z. (2009). Studies on the 5' exonuclease and endonuclease activity of CPSF73 in histone pre-mRNA processing. *Mol. Cell Biol.* 29, 21-32.

184. Ye, X., Franco, A.A., Santos, H., Nelson, D.M., Kaufman, P.D., and Adams, P.D. (2003). Defective S phase chromatin assembly causes DNA damage, activation of the S phase checkpoint, and S phase arrest. *Mol. Cell* 11, 341-351.
185. Yoshio, T.D., and Manley, J.L. (1997). RNA recognition by the human polyadenylation factor CstF. *Mol. Cell. Biol.* 17, 3907-3914.
186. Yuryev, A., Patturajan, M., Litington, Y., Joshi, R.V., Gentile, C., Gebara, M., and Corden, J.L. (1996). The C-terminal domain of the largest subunit of RNA polymerase II interacts with a novel set of serine/arginine-rich proteins. *Proc. Natl. Acad. Sci. U. S. A* 93, 6975-6980.
187. Zhang, Z., Fu, J., and Gilmour, D.S. (2005). CTD-dependent dismantling of the RNA polymerase II elongation complex by the pre-mRNA 3'-end processing factor, Pcf11. *Genes Dev.* 19, 1572-1580.
188. Zhang, Z., and Gilmour, D.S. (2006). Pcf11 is a termination factor in *Drosophila* that dismantles the elongation complex by bridging the CTD of RNA polymerase II to the nascent transcript. *Mol. Cell* 21, 65-74.
189. Zhao, J., Hyman, L., and Moore, C. (1999a). Formation of mRNA 3' ends in eukaryotes: Mechanism, regulation, and interrelationships with other steps in mRNA synthesis. *Microbiol. Mol. Biol. Rev.* 63, 405-445.
190. Zhao, J., Kessler, M., Helmling, S., O'Connor, J.P., and Moore, C. (1999b). Pta1, a component of yeast CFII, is required for both cleavage and poly(A) addition of mRNA precursor. *Mol. Cell. Biol.* 19, 7733-7740.
191. Zhao, J.Y., Kennedy, B.K., Lawrence, B.D., Barbie, D.A., Matera, A.G., Fletcher, J.A., and Harlow, E. (2000). NPAT links cyclin E-Cdk2 to the regulation of replication-dependent histone gene transcription. *Genes and Development* 14, 2283-2297.
192. Zhao, X., Kilop-Smith, S., and Muller, B. (2004). The human histone gene expression regulator HBP/SLBP is required for histone and DNA synthesis, cell cycle progression and cell proliferation in mitotic cells. *J. Cell Sci.* 117, 6043-6051.
193. Zhelkovsky, A., Tacahashi, Y., Nasser, T., He, X., Sterzer, U., Jensen, T.H., Domdey, H., and Moore, C. (2006). The role of the Brr5/Ysh1 C-terminal domain and its homolog Syc1 in mRNA 3'-end processing in *Saccharomyces cerevisiae*. *RNA*. 12, 435-445.
194. Zheng, L.-X., Dominski, Z., Yang, X., Elms, P., Raska, C.S., Borchers, C.H., and Marzluff, W.F. (2003). Phosphorylation of SLBP on two threonines triggers degradation of SLBP, the sole cell-cycle regulated factor

required for regulation of histone mRNA processing, at the end of S-phase. *Mol. Cell Biol.* 23, 1590-1601.